DOCUMENT RESUME

ED 147 367 TM 006 821

AUTHOR	Hedges, H. G.
TITLE	Achievement in Basic Skills. A Longitudinal Evaluation of Pupil Achievement in Language Arts and Mathematics.
INSTITUTION	Ontario Dept. of Education, Toronto.
PUB DATE	77
NOTE	363p.; Sections of the appendices may not reproduce well due to marginal legibility of original document
AVAILABLE FROM	Ontario Government Bookstore, 880 Bay Street, Toronto, Ontario M7A 1L2, Canada ($3.00); Ontario Institute for Studies in Education, Publications Sales, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada ($3.00)
EDRS PRICE	MF-$0.83 HC-$19.41 Plus Postage.
DESCRIPTORS	*Academic Achievement; Achievement Tests; Age Grade Placement; Arithmetic; Basic Skills; Comparative Analysis; Educational Change; Educational Objectives; *Elementary School Students; Elementary Secondary Education; Grade 5; Grade 6; Grade 7; Grade 8; *Language Arts; *Longitudinal Studies; *Mathematics; *Performance Factors; Predictor Variables; Reading Comprehension; Socioeconomic Status; Standardized Tests; Test Results; Vocabulary
IDENTIFIERS	*Ontario (Saint Catharines); Saint Catharines School System (Ontario)

ABSTRACT

A longitudinal evaluation of language arts and mathematics achievement for grades 5 to 8 in the St. Catharines, Ontario school system over a forty-year span is presented. The comparison is based on three different studies conducted from 1933 to 1938, from 1952 to 1954, and from 1975 to 1977. Secondary contributions--the effects of changes in relative socioeconomic status, age-grade levels, vocabulary, and school objectives--have been taken into consideration. The major conclusions are as follows: (1) Current grade 8 students achieve consistently lower than their earlier counterparts in arithmetic computation and reasoning; however, they perform as well as or marginally better in paragraph reading. (2) Today's grade 5 to 7 students outperform their earlier counterparts in fundamental operations in arithmetic. (3) The performance of grade 6 students in paragraph reading has shown a steady but small improvement over the past forty years. (4) The ability of students in grades 5 to 8 to initiate vocabulary from descriptions has sharply improved over their earlier counterparts, and reveals a much greater range of correct alternative responses. The arithmetic and language arts tests utilized in the 1975-77 studies are appended. (Author/MV)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED147367

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. A. Penny
Ministry of Education,
Ontario

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

© The Minister of Education, Ontario 1977
Queen's Park
Toronto, Ontario

Printed by The Ontario Institute for Studies in Education

MINISTRY OF EDUCATION, ONTARIO CATALOGUING IN PUBLICATION DATA

Hedges, H G
Achievement in basic skills : a longitudinal evaluation of principal achievement in language arts and mathematics / Principal Investigator : H. G. Hedges. -- Toronto : Minister of Education, 1977 (Ontario Institute for Studies in Education)

327 p.
"This research project was funded under contract by the Ministry of Education, Ontario."
Copies of tests bound as appendices.
Bibliography : p. 228-37. Appendices : p. 239-327
1. Academic achievement 2. Educational tests and measurements 3. Language arts (Elementary) - Examinations, questions, etc. 4. Mathematics - Examinations, questions, etc. I. Ontario. Ministry of Education. II. Title.
ON01101 371.262

Additional copies may be ordered from:

The Ontario Institute for Studies in Education
Publications Sales
252 Bloor Street West
Toronto, Ontario
M5S 1V6

or from:

The Ontario Government Bookstore
880 Bay Street
Toronto, Ontario
M7A 1L2

1 2 3 4 5 XX 08 97 87 77

Abstract

This report documents a study funded under contract from the Ontario Ministry of Education, designed to make a longitudinal evaluation of student achievement in language arts and mathematics in grades 5 to 8 in one school system over a forty-year span. The comparison is based on three major measurement studies conducted from 1933 to 1938 by Dr. Corbin A. Brown, from 1952 to 1954 by Dr. Hugh R. Partlow, and during 1975 to 1977 by the author. Although the central focus is a comparison of student achievement in the two subject areas concerned, the study also contains a number of secondary contributions which include consideration and adjustments for the effects of changes in relative socio-economic status, age-grade levels, vocabulary, and school objectives

The central objective of the study was to answer four main questions, as follows, about the populations under investigation:

1. How does pupil achievement in Language Arts and Mathematics in 1976 compare with results in 1938 and 1954, based on data from the same tests and schools?

2. How would a comparable group perform at present on the same tests? (The term "comparable" refers in this study to such factors as socio-economic background and age-grade patterns.)

3. How would a comparable group perform at present on a "fair" test? (In this instance "fair" refers mainly to characteristics of the tests reflecting changes in school objectives and vocabulary.)

4. What other factors may account for differences in achievement that remain after the foregoing factors have been taken into account?

In order to deal with these four questions and to provide a model and data for future longitudinal studies the study was designed according to following procedures and characteristics:

1. The same standardized tests were employed at the beginning, throughout, and at the end of the study;

2. All the students of the grades under study were included in all phases;

3. In each phase of the study the researcher employed comparable procedures for administering and scoring the tests;

4. Changes such as socio-economic variables, age-grade data, changes in program and objectives, and evolution of vocabulary were treated;

5. Where necessary because of sizeable changes in vocabulary and/or objectives, tests of fair comparison were added to the latest phase; and

6. Analyses of differences were undertaken not only in general or subject-wide achievement but also in major sub-components in each subject area under study.

The tests administered during the study were as follows:

1. Arithmetic Test of Fundamental Operations for Grade IV to VIII - Form A. (The Dominion Tests)
2. Test of Fundamental Operations for Grades 5 to 7: The Niagara Mathematics Test Series. (H. & M. Revision)
3. Group Achievement Test in Arithmetic Computation - Form A. (The Dominion Tests)
4. Group Achievement Test in Arithmetic Reasoning - Form A. (The Dominion Tests)
5. Group Achievement Test in Arithmetic Computation - Form B. (The Dominion Tests)
6. Group Achievement Test in Arithmetic Reasoning - Form B. (The Dominion Tests)

7. Test of Arithmetic Computation for Grade 8: The Niagara Mathematics Test Series. (H. & M. Revision)

8. Test of Arithmetic Reasoning for Grade 8: The Niagara Mathematics Test Series. (H. & M. Revision)

9. The Thorndike-McCall Reading Scale for the Understanding of Sentences - Form 2.

10. Group Achievement Test in Paragraph Reading - Form A. (The Dominion Tests)

11. Group Achievement Test in Paragraph Reading - Form B. (The Dominion Tests)

12. Vocabulary Test #1, 90 Common Names

13. Vocabulary Test #1, 90 Common Names. (Niagara Test Series)

Altogether, approximately 10 000 tests were administered, marked, and tabulated, in the latest phase, for students in the eleven study schools and in a few classes outside the study group where tests were administered for various comparisons.

The study is unique in Ontario's educational history; there appears to be no other system with a data bank of such duration based on identical or similar tests.

Of particular interest in this study are the strategies for the quantification of age-grade and socio-economic differences of students today with respect to their counterparts in the 1930's and the 1950's

The major conclusions, apart from those describing the effects of the four main variables, are as follows:

1. The grade 8 students do consistently worse than their earlier counterparts in arithmetic computation and reasoning.

2. The grade 5 to 7 students today, based on a test of fair comparison, outperform their earlier counterparts in fundamental operations in arithmetic.

3. The grade 8 students perform about as well as or marginally better than their counterparts in the two earlier phases in paragraph reading.

4. The performance of grade 6 students in paragraph reading has shown a steady but small improvement over the past forty years, as documented in the three phases of this study and an independent study conducted by Hastings in 1971.

5. The ability of students in grades 5 to 8 to initiate vocabulary from descriptions, as determined by a modern test of fair comparison, is sharply improved over their earlier counterparts, and reveals a much greater range of correct alternative responses.

These general conclusions are tempered and accounted for by more detailed considerations in the report. The reader is cautioned against making uncritical generalizations from this study to other populations without at least considering the similarities and differences in the systems being compared.

Contents

Abstract iii

Tables x

Figures vx

Preface xvii

1 BACKGROUND 1

2 DESIGN 8

3 INSTRUMENTATION 12

THE ARITHMETIC TESTS 13

The Dominion Arithmetic Test of Fundamental Operations, Grades IV to VIII, 1934 Edition, Form A 13

The Niagara Mathematics Test Series: Fundamental Operations, Grades 5 to 7 18

The Dominion Group Achievement Test: Part 2: Arithmetic Computation and Arithmetic Reasoning: Grade 8: 1934 Edition, Forms A and B 19

The Niagara Mathematics Test Series: Arithmetic Computation and Arithmetic Reasoning: Grade 8 22

THE LANGUAGE ARTS TESTS 24

Thorndike-McCall Reading Scale for the Understanding of Sentences, 1920 Edition, Form 2 24

The Dominion Group Achievement Test, Part 1, Paragraph Reading, Grade 8, 1934 Edition, Forms A and B 26

Vocabulary Test #1 (90 Common Names) 28

Niagara Test Series: Vocabulary Test: Revised Edition (1976): (90 Common Names) 35

SUMMARY OF INSTRUMENTATION AND ADMINISTRATION 40

4 CHANGES IN THE SCHOOL SYSTEM 44

ADMINISTRATION AND CONSULTATIVE SERVICES 48

SPECIAL EDUCATION SERVICES 50

GUIDELINES AND PROGRAMS.................................... 52

Format.. 53

Approach to Curriculum.................................. 54

General Aims.. 55

Content Expectations.................................... 56

Learning.. 57

Time-Tables... 58

Pupil Progress.. 60

Program of Studies: English............................. 61

Program of Studies: Mathematics......................... 61

THE EFFECTS OF LIVING AND LEARNING 62

SOCIO-ECONOMIC FACTORS.................................... 66

AGE-GRADE LEVELS.. 79

5 BACKGROUND TO THE DATA ANALYSIS.......................... 87

Rationale for the Study by Dr. C.A. Brown (1930's)...... 87

Rationale for the Study by Dr. H.R. Partlow (1950's).... 88

Instrumentation in the 1930's and 1950's................ 90

Familiarity of Staff and Students with Standardized Testing.. 92

Statistical Sophistication.............................. 93

6 COMPARATIVE ACHIEVEMENT IN MATHEMATICS..................... 96

GRADE 8 ACHIEVEMENT IN ARITHMETIC COMPUTATION AND REASONING.. 96

FUNDAMENTAL ARITHMETIC OPERATIONS, GRADES 5 TO 7........... 107

ITEM ANALYSIS: THE NIAGARA MATHEMATICS TEST SERIES, REVISED EDITION, FUNDAMENTAL OPERATIONS, 1976........... 125

7 COMPARATIVE ACHIEVEMENT IN LANGUAGE ARTS................... 142

GRADE 8 ACHIEVEMENT IN READING......................... 144

GRADE 6 ACHIEVEMENT IN READING......................... 151

COMPARATIVE ACHIEVEMENT IN VOCABULARY SKILLS, GRADES 5 TO 8... 159

SUMMARY OF LANGUAGE ARTS ACHIEVEMENT DATA................ 177

8 CONCLUSIONS 178
ACHIEVEMENT IN MATHEMATICS 178
ACHIEVEMENT IN LANGUAGE ARTS 188
THE TREATMENT OF MAJOR VARIABLES 195
EFFECTS OF OTHER VARIABLES 201
Class Size 202
Teacher Education 203
Teacher Autonomy 204
Classroom Methodology 204
Grouping Patterns 205
Remedial Teaching and Special Education 205
Teaching Materials 206
Experience in Test Writing 206
Time Allocation to Subjects 206
Increased Responsibilities of Schools 207
Supervision 207
Student Motivation 208
Homework 208
Television and Other Media 209
Philosophy and General Practice 209
RECOMMENDATIONS FOR FUTURE STUDIES 210
SUMMARY OF CONCLUSIONS 212
9 OTHER OUTCOMES OF THE STUDY 215
Response to Current Criticisms 215
Implications for Program 217
Local School System Interests 218
Age-Grade Levels 219
A Cooperative Approach 220
The Confidentiality Issue 222
Design For Longitudinal Studies 223
New Test Construction 224
Data Bank 225
Bibliography 228
Appendix A The Arithmetic Tests 240
Appendix B The Language Arts Tests 291

Tables

3:1 The Number of Questions for Each Type of Operation in the Dominion Arithmetic Test of Fundamental Operations, Form A, 1934 Edition 14

3:2 Validity of Items in the Dominion Arithmetic Test of Fundamental Operations for Grades 5, 6, and 7 in the Three Phases of the Study 16

3:3 Summary of Tests Used for Comparison of Student Achievement in Language Arts and Mathematics 1933-38, 1952-54, 1975-76 38

3:4 Classroom Enrolment Statistics for all Core Schools in the Longitudinal Evaluation of Mathematics and Language Arts Achievement, St. Catharines Public Schools: 1934, 1954, and 1975 39

4:1 A Descriptive Classification of the Socio-economic Status of School Populations in St. Catharines for the Period 1934-1954 and the Relative Socio-economic Status for 1975-76 for the Schools in the Study 69

4:2 Grade Means for Low, Average, and High Socio-economic Status Students of St. Catharines Public Schools for the Niagara Mathematics Test Series (Fundamental Operations), Grades V-VII, H. and M. Revision, 1976 72

4:3 Grade Means for Students of Low, Average, and High Socio-economic Status, for the Niagara Mathematics Test Series, Arithmetic Computation, Grade VIII, (H. and M. Revision), 1976 73

4:4 Grade Means for Students of Low, Average, and High Socio-economic Status, for the Niagara Mathematics Test Series, Arithmetic Reasoning, Grade VIII, (H. and M. Revision), 1976 73

4:5 Grade Means for Students of Low, Average, and High Socio-economic Status, Vocabulary Test of Common Names (Niagara Revision), 1976 74

4:6 Grade Means for Students of Low, Average, and High Socio-economic Status, for the Dominion Group Achievement Test, Part 1, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B, for 1975 75

4:7 Grade Means for Students of Low, Average, and High Socio-economic Status, for the Thorndike-McCall Reading Scale, Grade VI, 1920 Edition, Form 2, 1975 75

4:8 Approximate Adjustments in Students' Raw Scores on Selected Tests Attributable to Change in Socio-economic Status, 1954 to 1975-1976 77

4:9 Average Ages of Students for Grades 5, 6, 7, and 8 in the St. Catharines Public Schools of the Study, May 31, 1935, 1954, and 1975 80

4:10 Approximate Adjustments in Students' Raw Scores on Selected Tests Attributable to Changes in Mean Age-Grade, 1954 to 1975-76 86

6:1 Arithmetic Tests written by Grade 8 Students in the Three Phases of the Study 98

6:2 A Comparison of Achievement Scores in Arithmetic Computation and Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1952, 1975) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form A 99

6:3 A Comparison of Achievement Scores in Arithmetic Computation and Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1953, 1975) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form B 100

6:4 A Comparison of Achievement Scores in Arithmetic Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1952, 1975, 1976) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form A, and the Niagara Mathematics Test Series 102

6:5 A Comparison of Achievement Scores in Arithmetic Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1953, 1975, 1976) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form B, and the Niagara Mathematics Test Series 103

6:6 A Comparison of Achievement in Arithmetic Computation and Reasoning for Grade 8 Students Based on the Dominion Group Achievement Test (Part II, Arithmetic Computation and Reasoning, Grade VIII, 1934 Edition, Forms A and B) for 1934 and 1952 and the Niagara Mathematics Test Series (Arithmetic Computation and Reasoning, Grade VIII; H. and M. Revision, 1976), for 1976 105

6:7 Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), for the Dominion Arithmetic Test of Fundamental Operations, 1934 Edition, Form A, for Grade Five Classes 1936, 1954, 1975 108

6:8 Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), Dominion Arithmetic Test of Fundamental Operations, Grade 6, 1934 Edition, Form A, 1936, 1954, 1975 109

6:9 Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), Dominion Arithmetic Test of Fundamental Operations, Grade 7, 1934 Edition, Form A, 1936, 1954, 1975 110

6:10 Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 5 Classes: 1936, 1954, 1975 112

6:11 Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 6 Classes: 1936, 1954, 1975 113

6:12 Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 7 Classes: 1936, 1954, 1975 114

6:13 Percentage of Each Type of Question Missed: Dominion Arithmetic Test of Fundamental Operations, (80 Questions), All Classes, Grades 5, 6, 7: 1936, 1954, 1975 116

6:14 Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A, 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematics Test Series (Fundamental Operations) -- Grade 5 119

6:15 Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A, 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematic Test Series (Fundamental Operations) -- Grade 6 120

6:16 Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A, 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematics Test Series (Fundamental Operations) -- Grade 7 121

6:17 A Comparison of Achievement in Arithmetic Fundamental Operations for Grades 5, 6, 7 Students Based on the Dominion Arithmetic Test of Fundamental Operations (Grades IV to VIII, 1934 Edition, Form A) for 1936, 1954 and the Niagara Mathematics Test Series (Fundamental Operations), Grades V-VII, (H. & M. Revision), 1976 122

6:18 Niagara Mathematics Test Series, (Fundamental Operations), Grades V to VII (H. and M. Revision). Number of Correct Responses and Percentage of Incorrect Responses for each Question 133

6:19 Percentage of Students Missing Questions in Categories of Questions for the Niagara Mathematics Test Series, Fundamental Operations, Grades V-VII, (H. and M. Revision), 1976 137

7:1 Summary of Tests Used For Comparison of Student Achievement in Language Arts, 1933-1938, 1952-1954, 1975-1976 143

7:2 Ontario Norms (1934) and Summary Data for Grade 8 Classes, 1952, 1975 from the Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B 145

7:3 A Summary of Overall Means (With Adjustments for Socio-economic Status and Age-Grade, 1975) for Grade 8 Students, 1934, 1953 and 1975, from the Dominion Group Achievement Tests, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B 146

7:4 Class Means and Standard Deviations for Grade 8 Classes, 1975, from the Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B 147

7:5 Overall Distribution of Grade 6 Scores (November - 1938, 1953, 1971 and December - 1975) from the Thorndike-McCall Reading Scale, Form 2 150

7:6 Distribution of Grade 6 Scores from the Thorndike-McCall Reading Scale, Form 2, by Schools and Classes, (November - 1938, 1953 and December 1975) by Schools 154

7:7 Summary Data for Grade Six Classes, November - 1938, 1953, 1971 and December - 1975 (With Adjustments for the Age-Grade Differences in 1971, 1975), from the Thorndike-McCall Reading Scale, Form 2 157

7:8 A Comparison of Grade 5, 6, 7, 8 Student Performance, Based on Vocabulary Test Number 1, Ninety Common Names, 1933, 1953, and 1975 162

7:9 A Comparison of Class Medians, Vocabulary Test Number One, 1933 and 1953 and Class Means, Vocabulary Test of Common Names (Niagara Revision) 1976, by Grades 163

7:10 Class Means and Standard Deviations, by Grades, on the Vocabulary Test of Common Names (Niagara Revision) 1976 167

7:11 Overall Means and Standard Deviations, Grades 5, 6, 7, and 8, on the Vocabulary Test of Common Names (Niagara Revision) 1976 172

7:12 Comparison of Medians, Means, and Percentage Scores for Grades 5, 6, 7, and 8, from the Vocabulary Test Number 1, Ninety Common Names, for 1933, 1953, and 1975, and the Vocabulary Test of Common Names, (Niagara Revision) 1976 173

7:13 A Summary of Medians and Means (With Adjustments for Socio-economic Status and Age-Grade, 1976), for Grade 5, 6, 7, 8 Students, from the Vocabulary Test Number 1, Ninety Common Names, 1933, 1953 and the Vocabulary Test Number 1, Ninety Common Names (Niagara Revision), 1976 174

Figures

6:1 The Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976. Percentages of correct responses for each item at each grade level. 127

7:1 A Comparison of Grade Six Crude Score Means for St. Catharines Public Schools, 1938, 1953, 1971, and 1975 (With Adjustments for Socio-economic Status and Age-Grade for 1971 and 1975) Based on the Thorndike-McCall Reading Scale, 1920 Edition, Form 2 158

Preface

At a time when both educators and the general public are concerned to an unusual degree about evaluation and accountability, and many questions are asked of educators concerning student performance, studies of the kind reported in this document will provide the data from which rational judgments about student achievement can be made, and if necessary, plans set in motion to alleviate any discrepancies between actual levels of achievement and reasonable expectations for student performance. Although the study looks at only two areas, Language Arts and Mathematics, and is limited to only one group of schools in one large school system, it has the advantage of making comparisons over a forty-year period; for this reason it may stand out as a unique study of evaluation in Ontario's elementary schools.

At least four factors fortuitously came into focus at the same time, enabling this study to be done. The first was the public concern referred to above. The second was the availability of a research grant from the Ministry of Education. The third was the willingness of the Lincoln County Board of Education to have the study conducted in its jurisdiction. The last and most important factor was the availability of the data from the two earlier phases of the study, roughly in 1937 and 1954.

A study of this kind requires the assistance and cooperation of many individuals and institutions. Without the contractual agreement with the Ministry of Education the study could not have been conducted in such detail. The Lincoln County Board of Education, and particularly its Director, Mr. Roger Allan, were supportive of the study from the outset. Because the study required extensive testing in several rounds in both the central schools of the jurisdiction and in a sample of other schools, a great many principals, teachers, and students faced the problem of assisting with the data-gathering, often in addition to an already heavy testing program; we are pleased to acknowledge their generous assistance.

The work of administering, marking, tabulating, and analyzing the test results was carried on mainly by Lana Stahl, with periodic assistance by Al Hobbs, Nick Tymoszewicz, Joan Doyle, Joan Graziani, Terry Begg, Jean Hedges, Marjorie Hedges, and Allan Shynal. Four academics assisted us at critical phases of the study - Dr. L. A. Popp, now at the College of Education, Brock University; and Dr. F. G. Robinson, Dr. H. H. Russell, and Dr. J.R.B. Cassie, all of the Field Development Centres of O.I.S.E. The last named officer also assisted with the analyzing and writing of the report. Joan Doyle and Joan Graziani performed numerous technical functions, including the typing of this report.

A number of ministry officials, from both the central office and the Niagara Regional office, gave special assistance with one or more steps of the study, including evaluations of test instruments, selection of items for the revised tests, and identification and interpretation of changes in policy and practices.

The three researchers who had collected the earlier data, Dr. Corbin A. Brown, Dr. Hugh R. Partlow, and Donald Hastings, are still active in educational research and administration; they were highly constructive in their interest and support and generous with their data, reports, and advice.

Special acknowledgement is given to Mr. Allan Shynal, Research Officer at Niagara Centre of O.I.S.E., who accepted, throughout the study, the major responsibility for its administration, including hiring and supervising of staff, organizing schedules, revising tests, conducting interviews, analyzing data, and maintaining the heavy workflow involved.

It is anticipated that the results of the study will find three general uses. Despite the report's limits in scope and subject, it is hoped that its results will shed some light on current questions about student performance, in view of the fact that with reasonable caution the findings may be generalized to other school systems. Of even broader use will be some of the procedures used in the problem of comparing achievement from different periods of time, when matters such as socio-economic conditions, idiom, and school objectives, among others, have undergone changes which can be expected to have influenced student performance. Thirdly, the study will provide for researchers at some future date, perhaps a generation hence, a useful set of data complementing those which the two earlier researchers contributed to the present study. At such a future date, the current data should provide for an even more intensive and varied comparison than was possible at this time, because its complete file will record not only the valuable reduced data of the kinds that we inherited (tests, marking schemes, attendance and age-grade data, medians, etc.) but will also supply the tabulation sheets for every class on each test, with a large number of the test papers - in short, all the raw data needed to enable the researcher to make future comparative studies.

H. G. Hedges

1 Background

This report documents a study that was designed to make a longitudinal evaluation of student achievement in Language Arts and Mathematics in grades 5 to 8 of one school system over a forty-year span. The comparison is based on three major measurement studies conducted from 1933 to 1938 by Dr. Corbin A. Brown, from 1952 to 1954 by Dr. Hugh R. Partlow, and during 1975 to 1977 by the author. Although the central focus is a comparison of student achievement in the two subject areas concerned, the study contains a number of secondary contributions which will be discussed later in the report.

The almost complete absence of longitudinal studies of this kind in Ontario can be accounted for mainly by the absence of systematic records based on standardized tests over a long period of time. A second factor, no doubt, has been the large number of changes made in organizational and administrative structures during the past two generations.

The first set of data was an outcome of a system-wide program of evaluation conducted by Dr. Corbin A. Brown, who in the period in question, 1933 to 1938, was the inspector-administrator of St. Catharines Public Schools. These data provided a base-line for comparisons by Dr. Hugh R. Partlow in 1952 to 1954 when he served as an inspector of schools in the same city. For purposes of his comparisons over the period from the mid-thirties to the mid-fifties, Dr. Partlow used Ontario norms from the 1930's for a number of tests, thus supplementing the data from the Brown study for the earlier period. Thus both studies involved students in the same schools and the same jurisdiction; additionally, Ontario norms were used where appropriate. The second study was documented in a report entitled "A Comparison of St. Catharines Public School Standards in Arithmetic and Reading, 1933-1938 and 1952-1954", by Hugh Russell Partlow, which was submitted to the University of Toronto as his doctoral thesis. The study was later published under the title, Arithmetic and Reading, by Hugh R. Partlow (Toronto: Copp Clark

Co. Limited, 1955). In 1970-1971 Donald Edward Hastings, a principal in St. Catharines, conducted a follow-up study on one aspect of the Language Arts program, in grade six, and reported his findings in an unpublished graduate report in partial fulfillment of the requirements for the degree of Master of Science at the graduate school of Niagara University.

The existence and availability of these three earlier studies, coupled with Niagara Centre's long-term interest in studies of student achievement, led to discussions with officials of the Ministry of Education concerning the possibility of conducting a longitudinal study based on the earlier tests and results. When it was assured that approval and encouragement to conduct such a study would be given by the Lincoln Board of Education, in whose jurisdiction the St. Catharines schools in the earlier studies are located, the possibility of continuing with the study became a reality, and a contract was awarded by the Ministry of Education.

This study comes at a propitious time in Ontario's educational development. Following a decade of remarkable growth in the school system, coupled with major reforms in organization and administration, including the development of county school systems, greater flexibility and local autonomy in curriculum development, generous support for innovation and experimentation, the expansion of research and graduate instruction, and numerous other new structures and practices, there has emerged in the 1970's a trend toward closer scrutiny of educational objectives, organization, practice, and achievement.

The most obvious and persistent ingredient of the movement in the 1970's toward greater evaluation and accountability of student performance in schools has been the continuing public debate on the question among parents, educators, politicians, and the media. The debate is characterized on the one side by the claims of certain businessmen, university professors, politicians, groups of citizens, editors, and some educators that student performance, or standards of achievement, have fallen below what formerly existed; on the other hand many educational administrators, teachers, and citizens express concern about the procedures used by the critics in arriving at their conclusions. They state that the major changes in schools have been in philosophy, objectives, expectations, teaching practices, grouping and promotion patterns, time distribution among subjects, and the retention in school of greater percentages of the population, rather than in standards of achievement.

This widespread argument on the question of comparative student achievement is aggravated by the almost total absence of long-term evaluation studies based on standardized tests. Without such comparative data many of the critics and supporters of modern educational practices deal with highly personal, unsubstantiated comparisons of purported levels of achievement, the nature of the groups being compared, and the age or grade levels concerned. It is not surprising that difficulties of this kind occur in sincere attempts to make comparisons over a long period of time when much of the data relies more on memory rather than on documented records.

Although small, short-term, and informal comparisons of student achievement have been made, and while no doubt considerable amounts of unsystematic unanalyzed data exist in school board files, there appears to be an almost complete absence in Ontario of any studies of student achievement over a considerable period of time at the elementary school level which have applied the following essential characteristics of any study which would claim to make a systematic comparison of student achievement:

1. employs the same standardized tests at the beginning, throughout, and at the end of the study;
2. either compares all the students from the same schools or geographic area, or employs an adequate sampling device to ensure that proper matching of subjects is conducted;
3. employs in each phase of the study comparable procedures for administering and scoring the tests;
4. takes into account changes such as socio-economic variables, age-grade data, changes in program and objectives, changes in

vocabulary, etc., which could be expected to account for some of the differences in achievement over a period of time; and

5. analyzes differences not only in general or subject-wide achievement but also in major components of each subject area under study.

The purpose of the present study was to make a longitudinal evaluation of student achievement in two major subject areas, Language Arts and Mathematics, and in the process to meet the essential criteria cited above, and to apply other criteria which could be expected to make the outcome of the study useful for comparative purposes as well as for proposing curriculum emphasis or change.

As background to the present phase of the study, the major changes in student achievement documented by Partlow and Hastings following their respective studies should be noted. In comparing achievement in the two subject areas between the 1930's and 1950's, Partlow made the following major observations, among others:

1. the students in St. Catharines Public Schools made a significant overall improvement in Arithmetic and Reading from the 1930's to the 1950's;
2. in the seven sub-tests in Arithmetic all grades showed an improvement in the 1950's;
3. of the Reading Comprehension sub-tests, two showed an improvement in student performance and one showed no change in the 1950 phase of the study;
4. of the Vocabulary sub-tests, one showed an improvement in pupil performance, two showed no change, and one showed a decline in the 1950's;
5. despite the overall improvements, some declines were noted in certain areas in the two subjects concerned between the 1930's and the 1950's; for example, there was a decline at the grade 5 level in multiplication and division; in grades 5 and 7 in subtraction; and in multiplication in all grades tested; and

6. the greatest improvement in reading scores between the 1930's and the 1950's was shown by students in the upper levels of the classes.

The major conclusion of the Hastings study was that the improvements in reading noted by Partlow continued through the period between 1954 and 1971. It should be noted that the Hastings study looked only at the reading component in Language Arts and was conducted only at the grade 6 level. Hastings found that this general pattern of improvement was consistent in the 6th grade across the schools of St. Catharines; although the study was limited in scope, the researcher concluded from his own and the earlier studies that "continual pupil growth (i.e., improvement) in Reading Comprehension has been clearly shown over the past forty years."

With data and conclusions available from earlier studies using results with the same test instruments or normed data for Ontario, the present study was planned to repeat the process and compare the achievement levels today with those from the earlier periods. Therefore it was decided to test pupils in the same schools, at the same grade levels, in the same subject areas, and with the same tests. These conditions, then, would match those in the earlier studies.

In making a comparison of student achievement over a considerable period of time, reliable generalizations can be formed only if the researcher has access to similar populations and to tests which themselves do not put the students at any one period of time, it was decided that if comparative data from Because these conditions probably do not exist over an extended period of time, it was decided that if comparative data from different time periods were to have any real meaning, at least four factors which would favour the population at one or other period of time, or which would create differences in the populations, should be accounted for. The factors having such characteristics appeared to be the following:

1. possible changes in the socio-economic level of the school communities; reference will be made later in the study to

differences that can be attributed essentially to socio-economic factors;

2. age-grade levels of the pupils involved; it is obvious that any large change in the age distribution at various grade levels would have important bearings on the results;
3. changes in vocabulary; for those tests in which vocabulary was important, changes in the use of words over time would create disadvantages at later periods in a number of respects, including the effects of changes in popular use of words, changes in meaning, and problems in using a marking scheme from one period to assess answers at a later period;
4. changes in school objectives; it would seem unreasonable to expect students from one time period to achieve at the same level as those a generation earlier if major changes occurred in school objectives, in course content, time spent on each subject, emphasis on various subjects, etc.

In addition to the four factors stated above (socio-economic changes, age-grade differences, changes in vocabulary, and changes in school objectives), all of which must be taken into account in describing changes in achievement from one period to another, there are many other factors which can be expected to have had a positive or negative influence on comparative student achievement over a forty-year span. Although some of these possible causes of changes in student performance will be discussed in the report, no systematic attempt was made to measure separately the effects of each of the following possible factors:

1. changes in teacher education;
2. changes in time allocation to different subjects;
3. changes in classroom methodology;
4. changes in grouping patterns in classrooms or schools;
5. changes in amounts or type of supervision by consultants and other officials;
6. changes in teaching materials;
7. changes in motivation of students;
8. changes in the amount of remedial teaching;
9. changes in pupils' experiences in writing tests;

10. changes in the autonomy of the teacher;
11. changes in teacher-student ratio;
12. effects of television and other media;
13. changes in amounts or types of homework;
14. changes in educational philosophy; and
15. the addition or reduction of subjects and other responsibilities of the school.

While it is obvious that many and perhaps all of the factors listed would account positively or negatively to some degree for changes in student performance in schools, the specific directions and dimensions of such effects can be determined only as a result of further evaluation studies, some of which will need to isolate limited numbers of these factors and measure their impact. No doubt a few of these factors, positive or negative, are largely beyond the power of the school to influence, while others come under the direct or indirect influence of the school and therefore merit careful evaluation. Such systematic studies of other factors influencing student achievement, however, are beyond the intent and scope of the present project.

2 Design

The present study, dased on data on student performance from three periods, each about one generation apart (roughly 1938, 1954, 1976) compares achievement in Mathematics and Language Arts of the students in grades 5 to 8 of the group of schools in St. Catharines, Ontario, which prior to reorganization of school systems into larger administrative units represented the complete set of elementary public schools in that city. In most instances the data were from results on standardized tests administered and marked by trained and competent teachers and research officers, under conditions as comparable as possible; some of the data were from Ontario norms, and some from revised tests which, although carefully constructed and pre-tested, were not standardized.

The central objective of the study was to answer four main questions about the populations under investigation as follows:

1. How does pupil achievement in Language Arts and Mathematics in 1976 compare with results in 1938 and 1954, based on data from the same tests and schools?
2. How would a comparable group perform at present on the same tests? (The term "comparable" refers in this study to such factors as socio-economic background and age-grade patterns.)
3. How would a comparable group perform at present on a "fair" test? (In this instance "fair" refers mainly to characteristics of the tests reflecting changes in school objectives, and vocabulary.)
4. What other factors may account for differences in achievement that remain after the foregoing factors have been taken into account?

The purpose of this chapter is to describe the general design of the study. Details of the application of the design are dealt with in appropriate sections of the report. Procedures for administration are described in a subsequent chapter.

The design called for the procedures described below.

1. The administration of the original tests to the entire population in grades 5 to 8 of the schools involved in the earlier studies. Although some minor changes in school populations, school closings, etc. occurred, sufficient stability in identity and populations persisted to enable the comparisons to be made. The names of the schools appear in the three earlier studies and are retained on all our records, test envelopes, tests and tabulation sheets; however, throughout this report the identity of each school is shown by a number. This practice does not detract from the usefulness of the report for the general reader. Any reader, researcher, or official who needs to identify any of the schools may ask the author for this information.

2. A comparison of age-grade distributions in each school during each of the three phases of the study. This required a careful check of the ages recorded in official records, including registers, O.S.R. cards. and annual reports.

3. An estimate of the effects of changes in age-grade distributions.

4. A comparison of the relative socio-economic levels of the community served by each school during the three phases of the study. This comparison involved numerous interviews with city, provincial, and federal administrators, senior educators, retired teachers and principals, and sociologists; it also included comparisons of housing patterns and conditions, employment statistics, relative real estate taxation, mean salary levels, and immigration.

5. An estimate based on current research comparisons, of the effects of changes in socio-economic levels.

6. The administration of the same tests in a selected number of schools with current socio-economic conditions judged to be roughly the same as those of the schools in the study during its earlier two phases, in order to determine whether the

estimated differences suggested by the research literature were apparent.

7. An analysis of the tests in the light of their appropriateness with respect to modern school objectives and modern language. This activity involved four procedures:

a. a careful scrutiny of the tests by a number of referees, mainly consultants, principals, teachers and other persons identified as experts in the area in question, to identify test items related to objectives, topics, or content no longer considered important or relevant in general practice and test items employing vocabulary or meanings not commonly used in school or by school-age children;

b. a comparison of courses of study and textbooks from the earlier phases with those in use today, to check validity of items;

c. the administration of the original tests to a number of classes, mainly outside the population under study, followed by an analysis of responses to determine the validity of the criticisms made through the two procedures above; items were considered unsuitable if they were identified in the two types of scrutiny and obtained very low scores from a selected sample of students;

d. an examination of the acceptable answers on the scoring sheets, to determine whether the original acceptable responses were still appropriate, or whether changes in knowledge or language have created alternative correct answers.

8. The preparation of equivalent tests as substitutes where the amount of change in school objectives and language was excessive. The greatest amount of change in school objectives was found to be in mathematics programs; later in the report a description is given of the procedures used in preparing parallel tests reflecting modern mathematics objectives. Although changes in language were evident in all the tests, and also showed up both in the range of student responses and in the range of acceptable answers, the only test showing such a

high degree of change in language as to require major revision was the test dealing exclusively with vocabulary. In this case a new parallel test was prepared; details of the procedure employed in the preparation of this test are given later in the report.

9. An initial comparison of achievement in the three periods of the longitudinal study, without adjustments for the changes in the major variables of socio-economic factors, age-grade, school objectives, and idiom.

10. An interpretation of the data in the light of changes in the four variables, with a description of changes in performance after these variables have been taken into account.

11. An analysis of student performance on major components of the Mathematics and Language Arts programs.

12. An analysis of the apparent effects of additional factors to which the changes in performance not otherwise accounted for may be attributed.

13. Identification of the major findings of the study.

14. Identification and discussion of other important outcomes of the study, with implications for future studies.

3 Instrumentation

The purpose of this chapter is to describe and provide background information for all the tests employed in the study. Copies of all the tests are found in the appendix.

The essential criteria for selecting tests were: (1) that they appeared in one or both of the earlier phases (Brown 1938, or Partlow 1955), and (2) that in their original or revised form they were "fair" tests in the 1975-76 phase, that is, that they reflected current objectives and language to an extent which made them suitable for comparative purposes. It will be noted in the descriptions below that two of the earlier tests were judged to be "fair" tests today; that several of the tests required major amounts of alteration, resulting in the development of revised tests, for which the rationale and procedures are described; and that one test (the vocabulary test) was considered to be so outdated that a new parallel test using modern vocabulary had to be developed and administered. It should also be noted that in the case of the mathematics tests, both the original and the revised tests using current objectives were administered. The use of both the original and the revised tests enables one to judge the extent to which changes in apparent achievement are attributable to changes in objectives and course content.

In addition to revisions in the test instruments, a number of changes in the marking schemes were required for a variety of reasons. In some cases changes in the meaning of words in test items had become so pronounced as to cause a difference in correct or acceptable responses. In other cases the number of acceptable correct responses altered with changes in language. In a few instances the same test item today generates a quite different correct answer than a generation or two ago. The rationale for these changes is obvious - a student giving a correct response today should not be penalized for failing to give a different answer that was judged correct before changes in language occurred.

I THE ARITHMETIC TESTS

1. The Dominion Arithmetic Test of Fundamental Operations, Grades IV to VIII, 1934 Edition, Form A.

This test was administered to all Junior Third, Senior Third, and Junior Fourth (grades 5 to 7) students in March 1936, in March, 1954, and in March, 1975. The test is composed of 80 problems originally designed to range in difficulty from below grade 4 to the end of grade 8. The problems are largely mechanical in nature and include the four basic areas of addition, subtraction, multiplication, and division. These four fundamental areas are supplemented by problems dealing with weights and measures, fractions, percentage, and decimals. Working time for this test is 35 minutes. Allowing for the reading of instructions, the completing of the student identification section and the explanation of sample problems, the total administration time was 50 minutes. The numbers of questions and scores for each type of operation are summarized in Table 3:1.

In its original draft, the Dominion Arithmetic Test of Fundamental Operations was developed from a total of 480 test items. From this original pool of test items an experimental edition was constructed consisting of four similar forms. Each form consisted of 120 test items. A total of 5000 copies of each form was distributed to a test population of representative schools throughout Canada; all provinces were included in this original test population. Once this initial administration was completed, the test designers selected a pool of 160 of the most suitable questions from the original 480. From this selection two equivalent test forms, A and B, were constructed, each having a total of 80 items. The sections for addition, subtraction, multiplication, and division, weights and measures, addition and subtraction of fractions, multiplication and division of fractions, addition, subtraction, multiplication and division of decimals, and percentage were constructed with considerable care to match

TABLE 3:1

THE NUMBER OF QUESTIONS FOR EACH TYPE OF OPERATION IN THE DOMINION ARITHMETIC TEST OF FUNDAMENTAL OPERATIONS

FORM A   1934 EDITION.

Sections of the Test	No. of Questions (Each Question Has A Value of "I")	Types of Items
The Four Primary Operations	20	Addition (Group 1) Subtraction (Group 2 Multiplication (Group 3) Division (Group 4) 5 of each.
Weights and Measures	10	Short Problems with Weights and Measures (Group 5)
Fractions A	10	Addition and Subtraction of Fractions (Group 6)
Fractions B	10	Multiplication and Division of Fractions (Group 7)
Decimals A and B	10	Addition, Subtraction; 5; (Group 8) Multiplication, Division; 5; (Group 9)
Percentage	10	Short Problems (Group 10)
Decimals and Percentage	10	Short Problems (Group 11)
TOTAL	80	

sections in both Forms A and B as to level of difficulty and number of problems in each. In order to establish norms for the test in 1934, ten thousand students were tested throughout Canada with either Form A or Form B. The Ontario norms for the two forms were determined by using the test scores from students in the province.

In order to determine the reliability of this arithmetic test, the test designers used both the Spearman-Brown Prophecy Formula and the Odd versus Even technique. The results of this study indicated an overall rating of 0.97, with specific reliability ratings of 0.87 for grade 4, 0.87 for grade 5, 0.94 for grade 6, 0.92 for grade 7 and 0.93 for grade 8. In its time (1934) the test would have been considered a highly reliable instrument for all grades tested.

When Partlow attempted to determine a logical or circular validity of each of the test items in this arithmetic test, he inspected the prescribed courses of study for each grade in both 1936 and 1954, to determine whether or not each test item was on the course of study. His investigation showed that the test was valid to an extremely high degree in both periods, with slightly higher validity in 1954 than in 1936. For the 1936 period, only one mathematics textbook was authorized by the Department of Education. If an item appeared in the textbook for a specific grade level, it was considered as reasonably certain that the mathematical skills would be taught to solve that specific test item at that grade level. In 1954, several mathematics texts were approved by the Department for use in the schools of Ontario, and since they did not correspond entirely with the prescribed course of study, teachers were instructed to follow the course of study in their instruction of arithmetic fundamentals rather than follow the organization of the specific texts used.

Table 3:2 summarizes the findings for validity of the test in the three phases of the study. It should be noted that while there was very little change in validity in any of the grades in

the 1936 to 1954 period, there has developed a large change in the 1954 to 1976 period. During the 1976 phase of the study, the number of valid items was roughly one-half of the number in the earlier two phases at each grade level. These changes reflect the major shift in mathematics objectives and content during the period, from an earlier emphasis on computation and "corner store" skills, to the current emphasis on understanding and processes associated with "new math." Even in content areas that existed throughout the three phases, for example, multiplication and division, the nature of the problems in the tests are inconsistent with those youngsters experience today. An obvious example of this general type of difference is the number of digits in problems in computation items in the test; students today simply do not experience computation questions with the size of some numbers in the test.

TABLE 3:2 Validity of Items in The Dominion Arithmetic Test of Fundamental Operations for Grades 5, 6, and 7 in the Three Phases of the Study.

		Grade 5	Grade 6	Grade 7	Total
Total Valid Items (80 items/test)	1936	29	44	68	$\frac{141}{240}$
	1954	32	47	69	$\frac{148}{240}$
	1975	16	21	35	$\frac{72}{240}$
Percentage of Valid items in the test/grade	1936	36.25	55.00	85.00	58.75
	1954	40.00	58.75	86.25	61.67
	1975	20.00	26.25	43.75	30.00

In both the 1936 and the 1954 testing sessions the teachers administered and scored the tests. Time and budget constraints dictated flexibility in teacher administration and scoring, but in neither period were the teachers given advanced knowledge about the test to be given. Tester training with reference to administration and scoring was covered by the circulation of an information bulletin which outlined the details of the specific test administration procedures. Principals and teachers were told of the test period during Feburary of 1936 and 1954, and specific administration bulletins were distributed at the end of that month in both cases. In the current study the research staff of Niagara Centre administered and scored the test. This procedure was undertaken to minimize errors and variation in administration and scoring and to facilitate the collection of data in the schools with minimal interference with timetables. Since this test was the first of the group to be administered under the research contract, the area superintendents and the principals of the core schools were alerted initially by telephone, and then by written notification. These steps allowed for the airing of problems and the design and implementation of specific strategies to overcome them. Testers from O.I.S.E. were employed and trained through pre-testing of the instrument in schools outside the study and at all relevant grade levels. This allowed for the elimination of unique administrative problems and much closer supervision and quality control of administration during the actual test period. Scoring of the 1975 tests was closely scrutinized and similar procedures were used in training scorers to eliminate errors in scoring and compilation of the data. One research officer was given primary responsibility for the testing and tabulation of the data, in an attempt to ensure maximum accuracy and consistency in the marking, checking, and recording of test results.

2. The Niagara Mathematics Test Series: Fundamental Operations, Grades 5 to 7.

The periodic revision of the curriculum in mathematics during the past twenty-five years and the evidence on the low validity of the Dominion Arithmetic Test dictated the construction of a modern parallel for fundamental operations for grades 5 through 7. A group of six consultants was assembled to review the original tests used in the study and to design new tests to correspond in both form and difficulty to the earlier test and to be relevant to the current math curricula in the grades under study.

Because of the availability of extensive research and development projects in elementary mathematics during the past ten years, it was determined that tests developed and validated by Dr. D. Horn at O.I.S.E. and Dr. F. Morrison of the Ottawa Board of Education would be used as the resource base for constructing these mathematics tests. In addition to employing validity data from the O.I.S.E. and Ottawa Tests, the test constructors checked each item with current textbooks and courses of study. The reader is invited to examine the original test pools and concomittant validation studies of these two resource bases.

The design of this test followed the format of the original Dominion Test of Fundamental Operations. The test involves eighty questions ranging in difficulty from grade 4 to grade 8. The first forty questions correspond very closely to items of the original test and require students to record their answer clearly in the space provided for each problem. In questions in the second half of the test the students encounter a combination of items with multiple-choice answers and computation problems with a space provided for recording the answer.

This test was administered by teachers during March of 1976. As in the earlier version, a fifty-minute time period was used. The research staff from Niagara Centre functioned in a support role and provided both consultative and administrative assistance when required. All marking and data recording were conducted by the research staff under the responsibility of one research officer. All teachers administering the test were asked to record any unusual problems encountered in the administration of the test, especially in regards to the relevancy of test problems to their own math curriculum. The test population in the core schools was 1062 students.

3. The Dominion Group Achievement Test: Part 2: Arithmetic Computation and Arithmetic Reasoning: Grade 8: 1934 Edition, Forms A and B.

Both Form A and Form B of this test were administered to grade 8 students in the core schools in St. Catharines during 1975. The test was originally designed for use in Canadian schools at the grade 8 level, beginning in 1934. Since no official record was kept in regards to the administration of this test to St. Catharines grade 8 students during the period from 1934 to 1939, the 1934 urban school norms for Ontario were used as a basis for comparison.

The date of June 5th was selected during both 1952 and 1953 as the testing time for both grade 8 student populations in St. Catharines. Form A was administered during the 1952 setting at which time 357 students wrote this test. Form B was administered in 1953 at which time 355 students wrote the test. In 1975, both forms of the test were administered to relatively the same student population during the month of June. Conditions varied greatly during the 1975 testing period because it was near the end of the school year and many classes of grade 8 students were on field trips, and individual arrangements were made to complete the testing. The results of the testing show that in the 1975 administration 318 students wrote Form A of the Arithmetic Computation test whereas only 258 students wrote Form B of the same test. When the Arithmetic Reasoning Test is considered, 372 students wrote Form A whereas 307 students wrote Form B of the test.

The Arithmetic Computation sub-test is composed of twenty-five questions (problems) on each of Forms A and B. The problems range across addition, subtraction, multiplication, and division, and include questions involving fractions, decimals, and percentage. Standard working time for the test was maintained at eighteen minutes. Total test administration time, allowing for distribution of tests, completion of the student identification section, completion of the pre-test questions, answering of procedural inquiries, and collection of the tests amounted to thirty-five minutes.

The Arithmetic Reasoning sub-test contains twenty problems in each of Form A and Form B. The majority of the problems involve two or three steps. The topics covered by the problems in the test involved percentage, profit and loss, taxation, discount, linear measure, interest, area, volume, and sharing. The working time for the test was twenty-five minutes and total administration time averaged forty minutes. Both forms of the test were designed so that the least able students would make a score greater than zero, whereas a perfect score would be attained by only one student in a thousand. The scoring scheme was established to be purely objective, that is, the answer for each question could be either right or wrong.

The initial experimental edition of the test was created first by the selection of problems from several hundred suggested items. This edition was distributed throughout Canada, and once the 10 000 copies were returned and scored, the unsatisfactory items were discarded. The remaining pool of questions was matched in terms of difficulty and topic covered, and the resulting pool permitted the construction of both Forms A and B, in 1934. Canadian norms were established by the administration and scoring of 2800 tests distributed throughout seven provinces. The largest number of tests administered were in the province of Ontario and the highest percentage of these were administered in urban schools. As a result of the administration and scoring of the Ontario tests, norms were set which were considered highly reliable for grade 8. A format identical to the one used for assessing the Fundamental Operations Test was used for both Arithmetic Computation and Arithmetic Reasoning. This involved the Odd versus Even technique and the Spearman-Brown Prophecy Formula. The results of the two tests indicated that both Form A and Form B of the Arithmetic Computation Test had a reliability of 0.83, whereas the Arithmetic Reasoning test had a reliability of 0.67 for both forms. The empirical validity of each test item was established by examination through elaborate statistical analysis; this was achieved by the application of the upper-and-lower-thirds method of validation. The logical validity of each test item was established by examination of the grade 8 course of study for 1934, 1953, and 1975. The type of questions and wording corresponded closely to the 1934 textbook terminology whereas decreasing correspondence was evident in the study of the 1952-53 textbooks, and became even more pronounced in the study of 1975 texts. As a result of the reduced level of validity in 1975, it was deemed necessary to construct a parallel test in Arithmetic Computation and Arithmetic Reasoning at the grade 8 level. This dual test is discussed in the following section.

In the original 1934 norming process, the teachers administered the tests and the scoring was conducted by the Department of Educational Research, Ontario College of Education. In both 1952 and 1953, the teachers in the St. Catharines schools administered and scored the tests. Training for the administration and the scoring involved a meeting of both principals and teachers of grade 8 students on the afternoon before the testing schedule began. Test procedures were thoroughly discussed and the tests, along with circulars outlining the procedures, were distributed to the teachers and principals. No reference to "teaching to the test" was discussed in the two early phases of the study, but it was stated that the principals and teachers did not know in advance of the administrative meeting what subjects would be tested and what tests would be used. During the 1975 administration of the tests, the research staff of Niagara Centre administered and scored the tests without prior information given to the teachers.

The format described in the introduction of the Fundamental Operations Test was followed for the Arithmetic Computation and Arithmetic Reasoning tests. The test scoring and compilation of data was again the specific responsibility of one research officer. Each student's test was marked and double-checked to ensure the accuracy of the data collection process. The procedures and personnel employed ensured uniform administration and marking.

4. The Niagara Mathematics Test Series: Arithmetic Computation and Arithmetic Reasoning: Grade 8.

Both of these two sub-tests were constructed according to the format of the original Dominion Test. Teachers administered both sub-tests to grade 8 students in the core schools under study during the period from mid-May to mid-June of 1976. The research team functioned on an on-call basis to eliminate major problems with test administration. The test population involved 478 grade 8 students for each sub-test.

Arithmetic Computation: Grade 8

This test also employed the same format as the original Dominion Arithmetic Computation Test. It involved twenty-five individual problems ranging from columnar addition to division of decimals and simplification of square root. Questions #1 through #24 required students to compute and record the answer to each question in the space provided. Question #25 required the student to compute the answer in the space provided and to circle one of four answers provided.

As with the revised Fundamental Operations Test described earlier, test constructors selected items from the O.I.S.E. (D. Horn) and the Ottawa Board of Education (F. Morrison) pools of test items. All items were further checked against courses of study and textbooks in current use in the core schools.

Prior to testing in the core schools, the instrument was tested in other schools in the system. It was found that a time period of twenty-five minutes was required for completion of the test.

Arithmetic Reasoning: Grade 8

This sub-test involved twenty questions, which covered topics ranging from simple interest and international monetary exchange to Venn diagramming. In questions #1 through #12 the student was instructed to do any computation necessary and to record the answer in the blank space beneath each problem. For questions #13 through #20 the students were instructed to do any required computation in the space provided; when the solution to each problem was reached the student was required to circle clearly one of the four possible answers provided. The selection of topics relied, as in the computation sub-test, on the test pools from the O.I.S.E. and Ottawa Board of Education studies. Again, all items were checked by comparison with current textbooks and courses of study.

The sub-test required twenty five minutes for completion by the majority of students.

No major problems arose in regards to the administration, scoring, and compilation of results with either the computation or the reasoning sub-test. As in other tests in this study, the research staff took extreme care in scoring, totalling, and tabulating the test results. It is anticipated that all instruments of the Niagara Test Series will be validated further in the schools of the Niagara peninsula by June, 1979, and therefore should provide normative data for future studies.

II THE LANGUAGE ARTS TESTS

1. Thorndike - McCall Reading Scale for the Understanding of Sentences, 1920 Edition, Form 2.

This standardized test was used as an instrument for collecting data in both 1973 and 1975 and was initially used as a measurement device in 1938. Its form, construction, and content make it ideal for utilization in a longitudinal study. It has by far the highest validity of any of the original tests in the study. The test includes ten paragraphs and one poem. Groups of three or four questions follow each paragraph and the poem; altogether, thirty-five questions are involved in the test. The difficulty and complexity of the concepts in the paragraphs range from very easy to highly complex with the easiest encountered at the beginning and the most difficult encountered at the end. The test manual, 1931 revision, was used in the 1938, the 1953 and the 1975 administrations. Reading grade norms for the items range from grade 1.7 to grade 15.9. Equivalent T-scores and reading-age scores for each raw-score and grade-score were also computed by the test authors. Grade-scores were established by testing 88 000 students registered from grade 2 to grade 12. A total of 10 463 test cases were used in establishing reading-age scores. When the test was constructed in 1920, nine duplicate forms of the test were designed; only Form 2 was used in St. Catharines during all three test periods.

The working time for the test was thirty minutes with an additional fifteen minutes required for the distribution of tests, the reading of instructions, the completion of the student identification section, and the collection of the tests.

According to the manual of directions for this test, "the probable error of a pupil is 1.5 crude-score; 0.5 grade-score; 3.0 T-score; and 0.9 H-score. The probable error of 0.5 grade-score, for example, means that there is a one-to-one chance that a pupil's true grade-score does not differ more than 0.5 grade-scores from his obtained grade-score."

Logical validity of this reading scale was established by examining courses of study, the characteristics of the reading textbooks, and the Canadian Word List. These steps led to the reasonable conclusion that the vocabulary used in the paragraphs, the poem, and the questions were within the range of comprehension of the average grade 6 student in all three test periods, that is 1938, 1953, and 1975. Direct examination of the reading scale indicates that the initial paragraph is directed toward an upper grade 2 level of vocabulary whereas the last paragraph is the most difficult and is aimed directly at the grade 6 level. Inspection of the questions, however, indicates a great range of difficulty. The questions following the first paragraph can be correctly answered by grade 3 students, but errors occur with most adults, even those with backgrounds in graduate education when the last paragraph and the poem are attempted.

The frequency of certain errors in items that otherwise appeared easy or obvious indicated that the idiom used in these items is now outmoded, or that some other language factor not readily identified had a common negative effect. For example, many pupils did not understand the terms "ten minutes of seven," "gleaner," or the distinction between "giving away" and "selling." In one section the use of the girl's name Beauty, now extremely uncommon, confused many pupils, who interpreted the question "How many brothers had Beauty?" as if it referred to the characteristic of beauty rather than to a name. Despite the apparent decrease in validity of these and a few other items, they were not altered in the test.

In examining the books used as readers by the students in 1938, 1953, and 1975, it was judged that those in the latter two periods utilized a much higher percentage of the words in the grade 6 students' normal or daily vocabulary. The reading difficulty of the test was judged suitable in all three periods, but some of the concepts, and particularly those in the poem, were judged as too difficult in 1975, and probably were inappropriate in the earlier phases also. In fact, the concepts tested in the poem caused considerable controversy among teachers and graduate students to whom they were presented, with the general conclusion that the interpretation of the poem implied by the acceptable answers in the scoring key was probably inadequate if not inaccurate.

The procedures followed for administering and scoring this reading scale varied in all three periods under consideration. During the initial administration in 1938, the tests were administered and scored by the staff of each school, with teachers and principals as testers and scorers. In November, 1953, Dr. Partlow administered the test to all grade 6 classes in St. Catharines, and a secretary at the Board of Education was employed to assist Dr. Partlow in the scoring and checking of the tests. In November 1975, the tests were administered by the teachers in the schools; however, the research staff of Niagara Centre was on call to assist with administration. The procedure followed for scoring, checking, and recording the 1975 data was similar to that followed for the mathematics tests. The accuracy and consistency of the procedures was maintained to a high degree by having one research officer responsible for supervising the marking, checking, and re-checking of the test papers. In all three administrations the principals and teachers of the schools involved were unaware of the test contents until the actual test session. The 1975 administration involved a detailed instruction sheet for the teachers; any problems encountered prior to, during, or after the test were recorded by the tester involved.

2. The Dominion Group Achievement Test, Part I, Paragraph Reading, Grade 8, 1934 Edition, Forms A and B.

Since there are no specific comparative data for the 1933-39 period, the 1934 Ontario norms are compared to the mean scores achieved by grade 8 students in St. Catharines during the 1952, 1953, and 1975 administrations. When the Ontario norms were established, the majority of students tested were from urban schools; this fact therefore permits a direct comparison of scores for all three periods.

The comparison of St. Catharines scores with Ontario norms has many limitations, but it has at least one value. Comparing the achievement of grade 8 students during the 1952, 1953, and 1975 administrations can reveal whether there is a significant difference in the mean scores when compared to Ontario urban school norms of 1934.

The lack of data for the St. Catharines population in 1934, however, limits the comparison of the mean scores achieved by the test populations in 1952, 1953, and 1975, with those of St. Catharines grade 8 students in 1934. The larger number of drop-outs in the '30's and the more rigorous screening effect of promotion practices of the same period resulted in a somewhat select test population. A common practice of the time was to hold in grade 7 some students who probably would not pass entrance examinations at the end of grade 8; to what extent this practice was employed in St. Catharines is not agreed upon. These factors indicate, however, that the grade 8's of both the '50's and the '70's are a less select group than those of the '30's.

Each of the two forms of this test requires grade 8 students to read six paragraphs and then for each paragraph to select the most plausible answer for two, three, or four questions from sets of five alternative answers. Different reading components are tested, including detail, general significance, and inference, with each type included in each test. The questions range in difficulty so that grade 8 students of low ability can answer some, but only one in a thousand students can answer all correctly. The answers to all questions are objective.

In constructing the tests the authors selected the best items from hundreds used in a trial run. From the results of administration of these to 10 000 students, unsatisfactory items were eliminated. The Ontario norms, based on largely an urban sample, were established on data from 2000 students. Using the Odd versus Even technique and the Spearman-Brown Prophecy Formula, reliability was determined as 0.64 ± 0.04.

From comparison with courses of study and textbooks, the tests were regarded as being equally valid in 1934, 1953, and 1975. The vocabulary in the paragraphs was within the experience of students in each time period.

The time provided to complete the tests was eighteen minutes, with an additional ten minutes required for completion of the student identification section and review of the sample paragraph and questions.

An administrative format similar to the one used for the Dominion Group Achievement Tests in Mathematics was followed for the administration and scoring of these tests. Time and scheduling constraints required the teachers to administer the test during the month of June in 1952, 1953, and 1975. Each teacher was provided with a detailed instruction sheet. In 1975 the research staff of Niagara Centre was available for consultation and administrative duties if required. As in other tests, one research officer was given principal responsibility for the administration and scoring of the tests and a comprehensive checking system was established to ensure the accuracy of scoring and tabulating of the data.

3. Vocabulary Test #1 (90 Common Names)

The purpose of this test is to determine how successfully students can generate correct answers to descriptions of common objects or materials. There are ninety questions in the test. For most of the test, only one acceptable answer was given in the scoring key used in the 1930's. For sixteen of the items an alternative answer was accepted. In November, 1933, four similar vocabulary tests were given to students in grades 5 to 8. However, only one of these, the one herein discussed, was administered in 1953, when it was given to all grade 5 to 8 classes in St. Catharines public schools. The same test was given in 1976, and for reasons discussed below an alternative similar vocabulary test was developed and administered several months later to the same students.

The items in the test call for the provision of a suitable answer to the following twenty groups of definitions:

1. Parts of the Body
2. Trades
3. Rooms
4. Receptacles
5. Tools
6. Shopkeepers
7. Dwellings
8. Openings
9. Motor Cars
10. Headdress
11. Beds and Bedding
12. Assemblies
13. Flowers and Plants
14. Household Fittings
15. Books
16. Fruits and Seeds
17. Parts of a House
18. Materials
19. Divisions
20. Measuring and Testing

The origin of the test has not been clearly established. According to Partlow, it was originally prepared in England, and was not standardized. Little is known about its method of construction or its empirical validity at the time it was developed.

This test illustrates more profoundly than any of the other tests the problem of administering tests in which the vocabulary used in many of the items, both in the questions themselves and more particularly in the answers, has changed from one phase of the study to another. Brown noticed in 1933 that a few of the items were not appropriate to the Ontario scene, and he eliminated or substituted for items then considered inappropriate. In 1953, Partlow found that most of the items were as valid for 1953 pupils as for 1933 students, but that a few were not. He noted that certain items that were in common use in 1933 were seldom used in 1953 and that in some cases children in 1953 would use a different term to fit certain definitions. He identified the following eight items of these types, as examples of the change in vocabulary or common usage during the intervening years from 1933 to 1953:

1. Wood or metal box for keeping small supplies of coal in.
   (scuttle)
2. A window which stands out in front of the rest of the house.
   (bay window)
3. Metal or wooden box for keeping small quantities of tea in.
   (canister or caddy)
4. What do we call a man who keeps a store for the sale of such things as currants, rice, sugar, tea, coffee and cocoa?
   (grocer)
5. What do we call a man who keeps a store for the sale of cigarettes, matches, pipes and smokers' outfits?
   (tobacconist)
6. What do we call the portions of a bottle of medicine as we take them?
   (doses)
7. Large inn especially fitted up for travellers and holidaymakers.
   (hotel)
8. What do we call the covering which a woman sometimes wears on her head when doing the housework?
   (dust cap)

Partlow noted that in addition to these eight items, three others were seldom used in the vocabulary of the 1950's, namely "stationer," "hearth," and "baluster."

Partlow also found, when marking the papers using the 1933 scoring key, that it was reasonable to accept alternative answers for eight of the items and he adjusted the key accordingly. However, he found that the test itself and the scoring key in general were valid enough to justify their use, and therefore he found it unnecessary to alter the original test.

Since a fundamental criterion of the present study is that a "fair" test reflecting modern usage was to be used, this test was complemented by a revised parallel form, described below. On our initial examination of the original test, and on subsequent checks to determine the validity of items, a very large number of items were found to be invalid. The number of items which had disappeared from child vocabulary, as reported on by Partlow, obviously had expanded at a rapid rate in the period from the 1950's to the 1970's, to the point that on this basis alone fully one-third of items were no longer appropriate. Examples of definitions no longer in use, or employing inappropriate wording for today, include the following:

> What do we call a man who is paid to drive a motor car?
>
> What do we call a room at the top of the house just under the roof?
>
> What do we call a large open-work packing case for oranges, eggs, etc.?
>
> What do we call the tool for grasping and holding lumps of coal?
>
> What do we call a large blunt needle for drawing tape, ribbons, etc., through hems?
>
> What do we call a man who keeps a store for the sale of such articles as microscopes, spectacles, lenses, and opera glasses?
>
> What do we call a man who keeps a store for the sale of writing paper, pens, pencils, ink, etc., but no newspapers or stamps?
>
> What do we call the holes in a net or sieve?
>
> What do we call the rubber covers fitted around the rims of the wheels of a motor car?
>
> What do we call the hat which a policeman wears?
>
> What do we call the covering which a woman sometimes wears on her head when doing housework?
>
> What do we call the part of the bed which is usually stuffed with felt or kapok?
>
> What do we call the little wheels fitted to the legs of a bed to make it easy to move?
>
> What do we call the instrument used for testing whether a table surface is quite flat?

It should be emphasized that the concern in this analysis of the validity of the test items has to do not so much with their difficulty as with the change in the frequency of their usage either in the culture at large or in children's vocabulary.

A second serious problem with the test items and scoring key, more or less independent of the criticism concerning the reduction of usage in the culture, has to do with the range of acceptable answers. As stated above, only one acceptable answer for most items appeared in the 1933 scoring key. The number of items permitting alternative answers increased slightly in Partlow's marking of the tests. By 1975, not only were many of the correct responses from 1933 and 1953 no longer correct or no longer the best possible answers, but also, for a very large group of the items, a number of equally correct or acceptable answers were currently common in children's vocabulary.
One factor probably influencing this situation has been a general increase throughout the culture in the use of alternative words; another may be the inclination of teachers to accept a wider range of responses to questions in schools today; a third may be deliberate attempts by teachers to encourage students to vary their choice of words.

Some examples of questions generating a number of equally acceptable answers in 1975 for which only one or two responses were accepted in the earlier phases are the following:

> What do we call a room at the top of the house just under the roof?
>
> What do we call the tool used for cutting grass and clipping hedges?
>
> What do we call the tool used for leveling the surface of pieces of wood?
>
> What do we call the hat which a policeman wears?
>
> What do we call the covering which a woman sometimes wears on her head when doing housework?
>
> What do we call a number of people gathered together to go for a picnic or outing?

> What do we call a small woody tree-like plant or bush?
>
> What do we call the entrance passage sometimes found in a house?
>
> What do we call the metal from which the best knife blades are made?
>
> What do we call the instrument used for measuring how heavy things are?

In addition to the two general criticisms (lack of suitability of the test items in the light of changes in vocabulary and the restricted number of acceptable correct answers in the earlier phases of the study), there were a number of other aspects of various items which made them invalid. For example, the meanings of certain words in common use in both periods had changed to such a degree that their correct use today would have caused them to be marked as incorrect answers in the earlier two periods. An extreme example of such change in use is the word "dose," for which the common meaning today is quite different from its meaning as "units of medicine" in the two early phases of the study.

The principal investigator and the testers were not alone in their awareness of and concern about the apparent lack of validity of many of the test items. In almost every classroom where the test was administered, the pupils, and particularly the teachers were extremely critical of the test, and pointed out that many of the test items were outdated or unsuitable in other respects.

One other check on the validity of the test items was employed. The researchers administered the tests informally to a reasonably large number of adults, most of whom were able to answer numerous items which were unfamiliar to pupils. However, even among the adults many of the items were left unanswered, apparently because the correct responses have disappeared from common usage; frequently when the correct responses were given later to these adults they recalled having used the accepted word during their own childhood.

In order to check the initial criticisms of certain items on the test, and also to attempt to estimate the validity of items, the test was administered initially to a number of classes in schools outside the core group. The investigator's own set of apparently unsuitable items was not only confirmed but also expanded, because in these schools as many as one-third of the total number of items obtained scores of 0 to 5 in some or all of the classes tested. This fact indicated that either the definitions were no longer meaningful or appropriate, or that the pupils could not, or did not, generate what was accepted as the correct answer as indicated in the scoring key. As a result of these findings concerning the low level of validity of many test items, the following two courses of action were decided upon:

1. To administer the test despite these weaknesses, in order to make comparisons on the same test for the three phases of the study; and

2. To construct a parallel test of equal difficulty, but employing definitions and answers that were approximately as common in child vocabulary today as the items in the original test apparently were at the time of its construction. The procedure used in developing the parallel test is described in the section that follows.

The administration of the original test in 1976 was similar to that used in 1933 and in 1953. In all cases, the tests were administered by the teachers, following the distribution of a circular outlining procedures for principals and teachers. One change in the procedures between the earlier two phases and the 1976 phase was in the marking of the tests. In 1933 and 1953 pupils were allowed to mark the papers under the careful supervision of teachers. In 1953 teachers individually checked the students' marking. In 1975, the tests were marked entirely by the research staff of Niagara Centre, with one officer in overall charge to ensure uniformity of interpretation and marking. The number of students tested in the three periods were as follows: in 1933, 648 students; in 1953, 1729 students; and in 1975, 1672 students.

4. Niagara Test Series: Vocabulary Test: Revised Edition (1976): (90 Common Names)

In view of the investigators' findings concerning the low validity of this test and in order to determine the performance of students when a "fair" test, that is a test using modern language, was employed, it was necessary to revise the test completely, thereby developing a parallel test by selecting from a very large pool of replacement items of similar extent of usage and difficulty.

In order to assist with the construction of such a test and to ensure its validity with respect to modern child language, a team of 15 language arts consultants, grade 8 teachers, principals, and superintendents was assembled, all of whom were judged or recognized as having considerable background and competence in language programs at the senior elementary school level. Members of this team performed a number of contributing activities. Some of them analyzed the existing test in order to assess the range of acceptable responses for each item. All members of the team examined the original test for validity and pointed out items and responses that they believed to be unsuitable on the basis of their knowledge of child language and the vocabulary used in textbooks. Members of the team were then asked to produce alternative test items of usage and difficulty approximately equal to that apparently existing in the early phases of the study. Altogether, approximately 500 test definitions were supplied, each with the set of suitable correct responses. From this list, through consultation and pre-testing, the set of ninety items was selected for administration during the 1976 testing period.

Consistent with the format of the original test, the definitions were separated into the original twenty distinct groupings identified in the section above. Similarly, each group of definitions was composed of three to five questions. The task of preparing and selecting suitable test items, when the two criteria of similar levels of usage and difficulty were applied, proved to be extremely difficult, and during initial comparisons caused considerable disagreement among the referees. It was often difficult to discriminate between breadth of usage and level of difficulty, and these two criteria in most cases seemed to overlap to various degrees. Terms which for different reasons might be considered difficult, such as motorcycle, astronaut, etc., were judged to be easy because of their frequent use. In contrast, reasonably simple words were often judged to be difficult because of limited usage. Despite these difficulties, the referees were able finally to agree not only on what items to replace, but also on suitable replacements for them and the acceptable set of responses for each.

As with the original test, teachers were responsible for administering this test to students from grades 5 to 8. The research staff of Niagara Centre was on call to consult and/or administer the test depending on local school conditions. It is worth noting that in contrast to the original version, very few teachers had negative comments about this test, and many of them commented favourably on the suitability of the test items. As in the original test, all marking, checking, and tabulating was done by the research team under the supervision of one research officer. To match conditions during the administration of the original test, students were given fifty minutes to complete the revised test.

From a research viewpoint, a potentially serious problem was created by administering first the original test and later the revised version to the same population. In general, it is considered that relative to internal validity writing a test may affect later performance on the same or a parallel test. The eventual need for a complete revision of the test in question was not anticipated during the stage when the design of the study was translated into a timetable for administration of the original tests. As a result, this phase was completed before the revised test was available. Thus there was a possibility that students may have learned answers from the first test, and then retained and applied them on the second test. The situation that faced the investigator was a trade-off of the disadvantage of possible test effects on the one hand, and the advantages of using as many items as possible that were common to the two tests on the other hand.

On checking on this problem with a number of eminent researchers and evaluators, the investigator was assured that under the circumstances very little, if any, test effect would occur. First of all, about four months passed between the two test sessions. Also, the teachers did not "take up" the answers to the original test, and no teaching of the test items was likely to have occurred because the teachers did not retain or mark the tests, did not receive feedback on their students' scores, and did not see the marking scheme. In cases where pupils may have remembered and later inquired about items with which they had difficulty or found to be unusual, such learning activities would typically have occurred in relation to items eliminated from the revised test for these same reasons.

While, as stated, there probably was little test effect, it must be recognized that the procedure created a threat that should be anticipated and eliminated in similar longitudinal evaluations in which revised tests are used to study or account for effects of important changes between test phases.

There are a number of ways in which this problem can be avoided in future studies, provided that both versions of the test are available at the same time. One procedure is to reverse the order of giving the two versions in pairs of classes. A similar but superior process involves dividing each class into two halves, and reversing the order of administering the tests to the two groups.

TABLE 3:3

Summary of Tests Used For Comparison of Student Achievement in Language Arts and Mathematics 1933-1938, 1952-1954, 1975-1976. *Refers to Ontario Urban Norms

Name of Test	Grade(s) Tested	Date of Testing			Number of Pupils		
		1933-1938	1952-1954	1975-1976	1933-1938	1952-1954	1975-1976
1 Dominion Arithmetic Test of Fundamental Operations, Grades IV to Grades VIII, 1934 Edition, Form A.	5,6,7	March 1936	March 1954	May 1975	1277	1329	1063
2 Niagara Mathematics Test Series (Fundamental Operations), Grades V-VII (H. and M. Revision), 1976.	5,6,7	----	----	June 1976	----	----	1062
3 Dominion Group Achievement Test, Part II, Arithmetic Computation, Grade VIII, 1934 Edition, Form A.	8	June 1934	June 1952	June 1975	1000*	347	318
4 Dominion Group Achievement Test, Part II, Arithmetic Computation, Grade VIII, 1934 Edition, Form B.	8	June 1934	June 1953	June 1975	1000*	343	258
5 Niagara Mathematics Test Series, Arithmetic Computation, Grade VIII (H. and M. Revision), 1976.	8	----	----	June 1976	----	----	478
6 Dominion Group Achievement Test, Part II, Arithmetic Reasoning, Grade VIII, 1934 Edition, Form A.	8	June 1934	June 1952	June 1975	1000*	347	372
7 Dominion Group Achievement Test, Part II, Arithmetic Reasoning, Grade VIII, 1934 Edition, Form B.	8	June 1934	June 1953	June 1975	1000*	343	307
8 Niagara Mathematics Test Series, Arithmetic Reasoning, Grade VIII (H. and M. Revision), 1976.	8	----	----	June 1976	----	----	478
9 Thorndike-McCall Reading Scale, 1920 Edition, Form 2.	6	November 1938	November 1953	November-December 1975	370	441	280
10 Vocabulary Test Number I, Ninety Common Names	5,6,7,8	November 1933	November 1953	January 1976	1648	1729	1672
11 Vocabulary Test Number I, Ninety Common Names (Niagara Revision), 1976	5,6,7,8	----	----	April-May 1976	----	----	1573
12 Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Form A.	8	June 1934	June 1952	June 1975	1000*	346	425
13 Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition Form B.	8	June 1934	June 1953	June 1975	1000*	340	410
TOTALS					9295	5565	8696

TABLE 3:4

Classroom Enrolment Statistics for all Core Schools in the Longitudinal Evaluation of Mathematics and Language Arts Achievement, St. Catharines Public Schools: 1934, 1954, and 1975.

SCHOOL	1934			1954			1975		
	Grades	No. of Classes	Enrolment	Grades	No. of Classes	Enrolment	Grades	No. of Classes	Enrolment
Alexandra	K. - Sr.IV	17	706	K. - Gr.6	16	607	K. - Gr.6 & 2 Auxiliary	15	359
Connaught	K. - Sr.IV	8	354	K. - Gr.6 & 1 Auxiliary	23	820	K. - Gr.8 & 3 Auxiliary	20	476
Edith Cavell	K. - Jr.IV	8	304	Jr.Kg. - Gr.8 & 1 Auxiliary	16½	617	K. - Gr.6	17	393
Glen Ridge	K. - Sr.III	3½	116	K. - Gr.6	9	339	K. - Gr.6	8	212
Memorial	K. - Sr.IV	8	343	K. - Gr.6	11	439	K. - Gr.6 & 1 Auxiliary	13	293
Prince of Wales	K. - Jr.III	3½	144	K. - Gr.8	12	464	K. - Gr.8 & 3 Auxiliary	13	293
W.J. Robertson	II - Sr.IV & 2 Auxiliary	10	354	K. - Gr.6 & 1 Auxiliary	12	305	K. - Gr.8 & 3 Auxiliary	8	135
Victoria	K. - Sr.IV	12	515	K. - Gr.6	13½	490	K. - Gr.6	10	259
Queen Mary	Opened September 1, 1953			Gr. 7 & 8 & 5 Auxiliary	14	412	Gr. 7 & 8 & 3 Auxiliary	19	429
Westdale				Opened 1956			K. - Gr.8	17	485
Lady Churchill				Opened 1965			Gr. 7 & 8	12	327

An even better procedure involves working with pairs of classes and dividing each class in half to form four groups to be designated A, B, C, and D; in the initial administration one version of the test is administered to groups A and D, the other to B and C; then in the second administration the process is reversed.

SUMMARY OF INSTRUMENTATION AND ADMINISTRATION

Because of the number of tests, the length of descriptions above, and the similarity of a number of the tests, Tables 3:3 and 3:4 are provided to summarize the tests used, the schools in the study and their enrolments, the number of students tested in each grade, and the month of administration of each test.

From the foregoing descriptions it will be noted that the testing program differed in two important respects from the two earlier phases. The more important of these had to do with the construction or revisions of the tests themselves. This change has been discussed in respect to each of the tests and will be further treated in the chapters dealing with the analysis of results.

The second major change, namely the procedures for administering, marking,and re-checking the tests, merits attention at this time and will be treated further in the chapter dealing with other factors which may account for differences in test scores. As noted, most of the administration and marking of the tests in the current phase was conducted by the project team, whereas in the earlier two phases of the study the tests were generally administered and marked by classroom teachers.

In the current phase, counting both the schools under study, commonly referred to in this report as the core schools, and the set of other schools in which tests were conducted in order to make comparisons, a total of about 14 000 tests were written. Most of the tests were marked at least twice and a very large number of them were checked a third time to ensure uniformity of interpretation and marking. Thus the administrative team marked roughly 30 000 tests.

The administration of such a large number of tests, that great many classrooms enabled the administrative team to observe changes in testing conditions which should be identified, some of which may have had an influence on test scores. Some of these changes in testing conditions are merely introduced here and will be considered at greater length in appropriate sections of the report.

The time of year for administering many of the tests, namely is June, should be considered in any future phases of a similar longitudinal study. Weather conditions during the testing period were at times extremely hot and humid and no doubt had some effect on student performance. A more serious disadvantage in conducting a testing program in June is the competition from a large number of highly desirable programs, including planning for extended field trips, field days, out-of-school projects, end-of-year programs, etc., and a marked reduction in formal instructional or pencil-and-paper activities. In this general situation, motivation to write the tests and classroom discipline generally were less positive than one would consider ideal for a testing program.

A second issue of concern to the administration team which may have had an indirect effect on student performance on the tests, was the question of the attitudes of principals and teachers. Although, as stated earlier, excellent cooperation from school staffs was generally evident, nevertheless there were factors which resulted in negative attitudes on the part of a few principals and teachers; some of these attitudes may have transferred incidentally to students. There was no evidence of any deliberate attempts on the part of teachers or principals to affect the attitudes of students negatively, but subtle effects of such attitudes were evident in a few situations. Some of the factors influencing teacher attitudes negatively were the large number of tests administered, the time of year, the competition of the testing program with previously scheduled programs in the schools, the additional administrative load for principals, the nature of some of the tests themselves (a criticism earlier referred to in this chapter), the lack of feedback because of the confidential nature of the study, and the apparent feelings of threat or insecurity associated with tests administered by outside agencies despite attempts to forestall such fears.

A third area of concern had to do with students' preparations for test writing and certain factors within classrooms themselves. During the 1930's and 1950's students were more accustomed than they are today to writing formal examinations. The lack of formal test experience can be expected to have had negative effects on the test scores in the current phase of the study. The more relaxed atmosphere of modern classrooms also created a difference in the attention and concentration given to the tests. In many of the classrooms the generally informal atmosphere and the customary practice of permitting or encouraging discussion and the sharing of information carried over to an undesirable extent in the test writing, with the result that reasonably strict administrative practices had to be employed, which themselves may have contributed negatively to the atmosphere in which the students wrote the tests. In 1975 situations in which the teachers administered the tests it was obvious that a wide disparity existed in the instructions given to the students, despite the fact that memoranda to teachers were provided in an attempt to ensure uniform instructions. Finally, the nature of the furniture arrangements in classrooms today is significantly different from the situation in the 1930's and 1950's, when children normally worked at individual desks. The patterns of groupings about tables in many of the classrooms today aggravated the problems of administration of the tests, and no doubt produced both positive and negative effects on test scores which would be impossible to measure precisely.

These changes in administrative procedures have been identified for a number of reasons. First of all, while they can be expected to have had certain effects on student test scores, no design was built into the study to enable us to ascertain the precise amount of such effects. Secondly, the identification of these concerns may enable future researchers to reduce their effects.

The author believes that if these problems had been adequately anticipated prior to the administration of the test adjustments in timetabling, instructions, communication with teachers, arrangement of furniture, etc., could have been made to reduce undesirable effects which may have resulted. It is recommended therefore that any parallel study of this kind take into account the observations of the administrative team concerning such unanticipated changes in testing conditions.

The reader will note that the schools in the study are identified by name in Table 3:4. In the remainder of the report the schools are identified by numbers only. The numbers used are not in the same sequence as the names in the table. The decision not to identify the schools by name in the presentations and analyses of data was based on the belief that this practice had more benefits than disadvantages. In particular, this practice is intended to avoid the negative effects of circulating achievement measures for individual schools which might prefer that such information remain fairly confidential.

At the same time, the nature of the mandate from the Ministry and the local Board does not explicitly require that the foregoing condition be exercised. Therefore any educator, official, trustee, researcher, or other person demonstrating a legitimate interest in the identity of the schools in any of the tables or elsewhere in the report is guaranteed access to this information by the author.

4 Changes in the School System

The purpose of this chapter is to identify major changes in the school system, particularly as they relate to the schools in the study, with respect to organization and administration, supervisory assistance, curriculum, socio-economic conditions, age-grade levels, and other factors which could logically be considered to have had an effect on student performance during the period from 1954 to the present. As background to this description and analysis, a review of changes in similar factors in the period from 1934 to 1954 should be presented.

In his report in 1954, Partlow assigned two chapters to changes in administration, supervision, and organization, comparisons of the populations tested, and changes in teachers, program, and facilities. His descriptions of these changes are much more detailed than their counterparts in the present study.

A brief review of the major changes documented by Partlow for the period from 1934 to 1954 is given below.

Partlow stated that there were no marked changes in administration or policy during the period between 1934 and 1954. Aside from the addition of one county representative on the board of education, and the addition of a second senior administrative officer, the administrative structure remained unaltered during this period.

According to Partlow, the changes in supervision during the period were more marked, with the gradual addition of supervisors and consultants. By 1954 supervision and special services were provided in music, guidance, speech correction, remedial teaching, art, reading, and non-teaching time for principals. A program for home instruction was also instituted, and eight auxiliary classes were added to the system.

The organization of classes in schools was considerably altered during this period. The general addition of kindergarten, the introduction of the first junior kindergarten class, and reorganization into junior and senior schools followed a change in grade designation from books such as Junior 3rd, Senior 3rd, etc. to grades, ranging from kindergarten to grade eight. During the period, the average class enrolment was reduced from 40.4 to 36.0 pupils per class.

Policy and practice in promotion and year-end failure changed a great deal in the period. Partlow stated that in the 1934-36 period the highest rate of failure occurred in the grade 6 and 7 years; the screening effect for the grade 8 year was revealed by the very high figures in these two grades. By 1954, there was a drop of over 12% in non-promotions from grade 7 to grade 8. The percentage of students who were not promoted at some stage in their elementary school years dropped to an average of 6.1% in the 1952 to 1954 period, from its high of 18.2% in the 1934 to 1936 period.

The trend toward fewer non-promotions, reflecting greater attention to social promotion, was matched at the other end of the scale by a similar reduction in the amount of "skipping" or rapid acceleration. By 1954, only about 10% of the pupils completed the eight grades in seven years, whereas in the early 30's, a considerable amount of rapid acceleration was evident, with, as an extreme example, as many as 30 to 40% of the pupils taking Book Two (grades 3 and 4) in one year.

The elimination of the high school entrance examination at the end of the eighth grade, in 1949, coupled with more flexible arrangements for transferring students from the 8th to the 9th grade, resulted in an increase in the percentage of total enrolment gaining entrance to the secondary school. To illustrate this point, Partlow showed that the percentage of the grade 7 enrolment passing from grade 8 to grade 9 after the normal eight years increased from 66.3% in 1933-1935 to 86.4% in 1952-1954.

For these and other reasons, the percentage of pupils leaving school before the end of the eighth grade went down substantially, from 10.7% in 1933 to 1.6% in 1954.

Practices with respect to promotion and non-promotion, rapid acceleration, high school entrance standards, school drop-out patterns, etc., naturally produced changes in age-grade levels during the period. The average age of students in Senior Fourth (grade 8) in 1933 was 12.8 years. By 1953 this figure had increased to 13.8 years. During this period, the percentage of students considered under-age for their grade decreased from 20.3% to 6.4%, but the percentage of pupils considered over-age for their grade had decreased only from 28.9% to 20.9%. The difference in these two decreases accounts in part for the change in the average age of grade 8 students.

Partlow also conducted a survey of the changes in socio-economic status of the populations in the schools in question, and showed that in 1954 the increase in the number of students in lower socio-economic situations, relative to the rest of the school system, was twice as great as the increase in the number in higher socio-economic situations.

In his report, Partlow identifies a number of factors which resulted in marked improvement in the qualifications of teachers during the period and in the opportunities that teachers had to improve themselves further.

The most important change that took place with respect to the curriculum during the period from 1934 to 1954 was the introduction of a new program of studies for the province, in 1937 in the first six grades and a year later in grades 7 and 8. The introduction of the new program was intended to bring about changes both in philosophy and in content. Prior to 1937, the program was prescriptive, achievement levels were relatively uniform and standardized, and the aim of the school was to make children fit into the uniform standards; emphasis was on mastery of subject material, with little attention given to individual interests and needs.

In contrast, the new programs of studies focused on the needs of the individual pupil, with attention given to individual differences in mental ability, pace, and interests. Following the introduction of the new program, further instructions and interpretations were provided both provincially and locally to teachers. Partlow states that the new programs had a great impact on educational methods and were in general well received and successfully implemented by teachers in St. Catharines. In his report, he described in some detail the specific changes in course content and methods related to mathematics and reading. Some of these changes were that less time was spent on drill in arithmetic; less emphasis was given to speed tests; there was an increase in silent reading; there was a trend toward grouping within the class according to the abilities of pupils; and there was an increase in problem-solving activities in arithmetic. One of the more controversial changes was the reduction in phonics in primary reading programs. Some of the more general changes in school practice were increased attention to motivation; competition with self rather than with classmates; a change from marks to letter grades on report cards; less regimentation in most phases of the program; a reduction in homework; an increase in the use of audio-visual aids; a wider range of approved textbooks; the introduction of new readers; and the addition of or improved provisions for such subjects as science, art, music, physical education, and religion.

Most of the major trends identified by Partlow in the period from 1934 to 1954 continued at varying rates, sometimes modest but in other cases accelerated, during the period from 1954 to 1976. Because many of these continuing trends have created alterations in school programs which logically could be reflected in achievement scores, it is the intention of the section below to identify changes in such areas as administration, socio-economic conditions, age-grade levels, special services, and the curriculum. It should be noted that three of these areas - socio-economic conditions, age-grade distribution, and changes in the curriculum - are three of the four variables identified in the design of the current project. Therefore, reference to these three variables will also be made in the chapters that follow. A detailed analysis of changes in the fourth variable, vocabulary, is beyond the limitations of the present study, and will be considered only to the extent that changing vocabulary was considered as a factor in changes in student performance on various tests.

ADMINISTRATION AND CONSULTATIVE SERVICES

The most obvious and in some respects the most far-reaching of the changes in the school system that occurred during the period from 1954 to the present was the establishment of county boards of education. Throughout the first half of this period, there was a gradual reduction in the number of school boards in the province, as neighbouring school systems became amalgamated.

In 1962, two changes took place in the St. Catharines public school system. At that time the schools in three surrounding municipalities, Grantham, Merritton, and Port Dalhousie, were amalgamated with St. Catharines to form an enlarged board of education. In the same year, a supervisory officer was appointed to look after secondary schools in the enlarged jurisdiction; this appointment represented a departure from the previous practice of having supervisors for elementary schools only.

The next, and by far the most significant of the changes in administrative structure that have occurred in the jurisdiction in several generations was the development of the Lincoln County Board of Education, following provincial legislation establishing county boards. This process occurred in 1968 and 1969. With the inclusion of all publicly supported elementary and secondary schools in the county under one board of education, with roughly proportionate representation of trustees from the various municipalities in the county, the small group of schools in the heart of the current study, which had been the entire set of elementary public schools in the system in the early two phases of the study, became only a relatively small proportion of the total number of elementary schools. At the same time, the program of the new board became an integrated kindergarten to grade 13 system in all areas.

In addition to the revised administrative structure that resulted from the development of county school boards, a number of other important changes in the local school system took place. It is considered that the most important general changes in school systems in the decade since the establishment of county boards have been provisions for equality of opportunity for students across the many types of municipalities in each system, the standardization of policies across the many schools in each system, the continuing improvement in teacher education (only part of which can be attributed to the development of larger boards), and a vast increase in the amount of curriculum development conducted in local schools and school systems. Senior officials of the Lincoln County Board of Education agree that these four general improvements are also characteristic of the local school system. The set of schools in the present study, the public elementary schools in St. Catharines prior to amalgamation and the establishing of a county board, shared in most of the major changes and improvements in the school system resulting from these administrative changes. However, in one respect at least the move toward equalization of opportunity can be considered to have had a negative effect, at least when compared to improvements made in the rest of the enlarged jurisdiction. This particular factor has to do with the amount of supervisory and consultative service available to individual schools in the former city of St. Catharines. In the period immediately prior to 1968, the school systems in Lincoln County had twelve full-time supervisory officers, four part-time officers, and the full-time equivalent of twenty-five consultants. In 1976, the total supervisory and consultative staff consisted of eight full-time supervisory officers and twenty-one consultants, including those in special education. Thus it is clear that a smaller administrative and consultative force has had to spread its services over about the same number of schools. It follows that the amount of administrative and consultative service to each of the schools in the present study has been somewhat reduced. To the degree that administrative and consultative support of teachers enhances professional development and the quality of instruction, the reduction in these supplementary services can be expected to have had an important negative effect on student performance in recent years.

SPECIAL EDUCATION SERVICES

One of the most obvious changes to occur in organization in the Lincoln Board has been a marked expansion in special education services. Changes have occurred in the number of personnel involved, the type of service, and the scope of activities. The policy statement describing provisions for special education in the system are included in the Special Education Services Manual, 1975-76

> It is the policy of this board to provide special education to help ensure that all children, whether they be normal, gifted, or handicapped have the maximum opportunity for successful school experience. It is also the policy of this board that special education students should be integrated into the regular school program as completely as possible according to their individual needs.

The staff of the central office of the Lincoln Board responsible for the administration of special education services consists of an associate superintendent, a consultant in charge of special education classes and programs, an attendance counsellor, a psychometrist, and various support personnel. These officers, along with other school personnel, including special education teachers, in cooperation with parents and community agencies, provide a wide range of special education identification and instruction programs. Included in these provisions are opportunity classes, remedial classes, special learning disability classes, a centre for children with emotional and learning disorders, pre-school and school age hearing impaired programs, classes for trainable retarded children, a crippled children's centre, major achievement programs for the intellectually gifted, programs for newcomers to Canada, remedial resource centres, an itinerant remedial language department, home instruction services, orthopaedic transportation units, limited vision units, and a special education services volunteer program.

The improvement in special education services in recent years has special significance for the present study, in that it has produced changes in the heterogeneity of the classes under study, and certain changes in the age-grade statistics. Since the expansion of special services, many students who would not typically have remained in school are now receiving remediation either in special classes or in regular classrooms. Many of these students are eventually integrated into the regular stream of the elementary schools. Thus many students who would have been segregated at other times, or who would have become drop-outs and therefore would not appear in the data, are retained and provided for in regular classrooms, including those in the study. This improvement in programs for children with learning difficulties represents a more humane, compassionate, and effective approach to youngsters with learning difficulties, but for the purpose of the present study has no doubt had a natural negative impact on test scores.

Both the provision of special classes for students with learning difficulties and the complementary practice of integrating as many of these students as possible into regular classrooms have a special bearing on the measures of central tendency used in the three phases of the study. Both Dr. Brown and Dr. Partlow used medians to represent typical class performance on the tests. In the absence of data concerning means, which would have been useful to the present investigator, the use of medians by the original authors brings with it certain fortuitous statistical advantages. For example, the use of medians in the two earlier phases of the study results in statistics that were unaffected by the size of extreme scores; this fact suggests that the very low and the very high academic achievers who would show up in the original data because no other special allocations were made for them are not taken into account in the measures of central tendency in the two earlier studies.

As stated above, since the expansion of special education services students at the ends of the ability range have remained in school programs and are found both in regular classrooms and in special education classes. In cases where students at the extreme ends of the continuum are removed from the regular elementary stream and given special instruction, the use of means in the present study gains rational support. While the means would have been altered by the extremes in the original studies, their presence did not affect the measure of central tendency because of the use of medians. Similarly, extreme cases today who are provided for in special classes do not participate in the testing program, and therefore the use of means in the current phase appears to provide a reasonable comparison. To give additional clarity to current measures of student achievement, standard deviations accompany the use of means. Two other groups of students who are in the classes tested, however, do affect the data today, and in both cases almost certainly have negative effects on mean scores. One group consists of students who would have left school had they lived in earlier phases of the study, particularly in the 1930's, but who remain in classrooms today as a result of special provisions made for them. The second group, already referred to, are those students who are recognized and provided for as special education students, but who are integrated into regular classrooms and therefore contributed their test results to the latest phase of the study.

GUIDELINES AND PROGRAMS

The following comparative account of guidelines, curricula, and programs in Ontario schools between the years 1937 and 1975 was completed with the generous cooperation of the Niagara Regional Office of the Ontario Ministry of Education. These general descriptions of changes are as applicable to the schools of St. Catharines as to most other jurisdictions. The details are considered briefly in nine categories.

(a) Format. As has been noted above, prior to 1955 the philosophy, suggestions, and course outlines by grade were presented in single documents from the Ministry, the more important of which were the Programme of Studies for Grades One to Six (1937), commonly entitled "the grey book," and its counterpart, "the blue book," for grades 7 and 8 (1938).

The program of studies for the first six grades remained almost unaltered until the mid-sixties. On the other hand, the documents for grades 7 and 8 faced periodic alterations. In 1949 the Department of Education announced the organization of school programs in four divisions - Primary, Junior, Intermediate, and Senior, with grades 7, 8, 9, and 10 combined for curriculum purposes into one division, the Intermediate. During the following year school systems were encouraged to commence curriculum revision and development for this division. A year later, the Department issued the Intermediate Division Outlines of Courses for Experimental Use, 1951. Then followed, for more than ten years, periodic memoranda and detailed program outlines for this division, each presenting the Ministry's policies and course of study in separate subjects.

After 1967, the practice of providing separate curriculum documents at different times was also applied to the Primary and Junior Divisions.

During the period from 1971 to 1974, the Ministry conducted a cyclical review of guidelines at the primary and junior levels and commenced a similar exercise for the intermediate level. In response to feedback received during the validation stage, the Ministry in 1975 divided the primary-junior guidelines into three separate parts which included policy, goals, and practical suggestions. During 1975-76, the documents distributed to schools included The Formative Years, which outlined Ministry policy for the primary and junior divisions, Education in the Primary and

Junior Divisions, which presented the philosophy and rationale behind The Formative Years, and a series of support documents presenting educators with practical classroom suggestions by which they might facilitate the implementation of Ministry policy.

Although surveys indicate that the documents distributed in the past several years are now reasonably well known by school staffs, their recency also suggests that they have as yet had relatively little impact on school programs, and therefore on the test scores obtained in 1975 and 1976. For this reason, it is more important to consider changes that occurred in the years between 1954 and 1975.

(b) Approach to Curriculum. As already noted above, Partlow identified changes in philosophy and provincial policy after 1937 which provided for greater measures of local and classroom autonomy and interpretation than had formerly existed. In comparison with contemporary practice, however, the approach to curriculum policy throughout the period from 1937 to 1955, and for some years thereafter, retained and reflected considerable amounts of curriculum control by the Department of Education. Policy was set by the Ministry, and its inspectors interpreted policy to schools. The administrative and consultative services to most boards were provided by ministry officials. Although these provisions suggest a large measure of central domination or control, the actual conduct of ministry officials generally was in the direction of encouraging teachers to be more autonomous, active, and responsible in providing for the needs of students in their classes and of making decisions at local and school levels. Thus by 1967 there was in fact a very wide variety of programs, practices, and interpretations of centrally generated policies. Between 1967 and 1971 the Ministry of Education moved formally from the provincially directed approach. Boards of education were deliberately encouraged to develop their own curricula, preferably at the school level, based on ministry guidelines. As a result, between 1971 and 1975 a great deal of curriculum development at local and school levels occurred in almost every system, and many excellent curriculum development projects and courses flourished with the encouragement of the Ministry. By 1975 the Ministry had moved from encouragement to the position that it was the school system's responsibility to participate in curriculum development. This general movement from provincial authority to local direction continued to the point that in 1975, broad ministry guidelines were used mainly as a basis for local board curriculum development, at least at the primary and junior levels. Ministry policy in this respect stated that:

> Thus while the Ministry articulates the broad goals, it is the responsibility of the local school boards...to formulate local programs that are within the rationale of the provincial policy and at the same time reflect local needs in priorities. School staffs both as individual teachers and as a collective body under the leadership of the principal, have the task of planning curriculum programs specifically adapted to the children for whom they are responsible.
>
> Those most aware of the children's needs and the community's expectations - parents, teachers, principals, supervisory officials, as well as the children themselves - must all be involved in the planning process in appropriate ways.

(c) General Aims. The main differences regarding the aims of education as presented in guidelines for the years 1937, 1941, 1955, 1967-1971, and 1975 are in specificity and intention. The aims as expressed in Education for Democratic Living, a Department of Education document for the 1931-1936 period, were retained in the same form throughout the first two phases of the present study:

> The schools of Ontario exist for the purpose of preparing children to live in a democratic society which bases its way of life upon the Christian ideal.

In the years from 1967 to 1971 the preceding quotation was retained as the broad aim of education. Added to it, however, was a list of specific characteristics of the child in the democratic society; this list included such concerns as self-realization, security, happiness, justice, opportunity to participate in all decisions affecting him or her, and so forth.

In 1975 the Ontario Government's position on curriculum policy, as presented in The Formative Years, stated that the curriculum, the planning and implementation of which were local responsibilities, was to provide opportunities for each child to grow to the limit of her or his potential.

> To acquire the basic skills fundamental to his or her continuing education;
>
> To develop confidence and a sense of self-worth;
>
> To gain the knowledge and acquire the attitudes that he or she needs for active participation in Canadian Society; and
>
> To develop a moral and aesthetic sensitivity necessary for a complete and responsible life.

Taking into account the multi-cultural context of Canadian society and predominant social trends, the Ministry made two noticeable changes in the 1975 statement of aims of education. There was a new emphasis on moral and aesthetic sensitivity, and a statement about provision against sex-role stereotyping.

The list of identified local board responsibilities has also expanded in the documents in recent years, to include not only local planning of curricula, but also individual diagnosis and remediation, the provision of space requirements, and parental involvement.

(d) Content Expectations. During the period between 1937 and 1976, expectations were stated mainly in terms of content for each grade level. The amount of specificity of content varied from subject to subject, with well defined sequences in the skills areas, considerable option in science and social studies, and much more freedom of choice in music and art. The encouragement of programs to meet individual needs, coupled with improvements in teaching skills and teacher autonomy, has resulted in an increasing range and variety of expectations. More and more attention was given to processes and attitudes as well as to the mastery of content. Projects, enterprises, and other individual and group activities were developed to reflect a broad range of objectives, including social and problem-solving skills. By the end of this period expectations, programs, and instructional practices began to reflect broad theme approaches, integrated units, and cross-disciplinary studies. In many respects the lines between different subject areas had become blurred. Expectations were stated in terms not only of traditional content, but also of physical, social, and mental skills, attitudes, creative expression, and individual traits.

In Ministry documents developed for the primary and junior divisions during the period from 1967 to 1971, recognition was given to broader, integrated approaches such as those identified above. However, with the shift away from specific content and the recognition of wider sets of expectations, concern was expressed that adequate planning should take place, and emphasis was placed on planning over longer periods of time.

> A practical efficient time-table requires careful long- and short-range planning...considering each day, each week, each month, each year.

Prior to and during this period considerable attention was given in the literature, and in pre-service and in-service programs, to behavioural objectives as expectations in school programs. This trend was recognized in 1975 in The Formative Years, which defines expectations in terms of student competencies.

(e) Learning. As could be expected, ministry guidelines from 1937 to the present have expressed references to learning more and more in terms of the psychological theories by which educators might respond to school objectives. The 1937, 1941, and 1955 guidelines stressed that "programs should be planned around the nature of the development of a child and the nature of the learning process." The educator of the 1960's and 1970's, unlike his predecessors, had available the works of Piaget, Bruner, Gagné, Ausubel, and other psychologists and educational theorists. It is not surprising then to note that the guidelines for the 1967 to 1971 period pointed out that "activities are to be functional, real, and appropriate for the child's level of understanding." Such statements recognized not only advances in learning theory but also improvements in teacher education programs, and implied not only that resources were at hand to enable educators to create more sequential learning experiences for youngsters, based on current knowledge of the psychological development of children, but that modern educators themselves could seek out theoretical positions from which to form judgments.

(f) Time-Tables. The guidelines of 1937, 1941, and 1955 gave general recommendations for the distribution of time to various subjects, but also encouraged flexible use of these recommendations. The general suggestions were as follows:

English	30%
Social Studies	20%
Health and Physical Education	10%
Natural Science	10%
Arithmetic	10%
Music	10%
Art	10%

In applying this time-table, 10% means roughly one-half hour per day.

At the beginning of this period, in 1937, many principals and teachers regarded this proposed distribution of time as being mandatory and made careful attempts to provide for these time allocations through isolated subject blocks. Recent interviews with a number of retired and senior teaching staff members indicate that their impression was that while there was a general encouragement of greater flexibility in timetable provisions throughout this period, many teachers continued to follow these suggestions rigorously. For the majority of teachers, however, at least two kinds of major departures from these timetable suggestions occurred. First, teachers began to feel more and more autonomous with respect to adjusting the allocation of time. The development of class projects, field trips, integrated units, theme approaches, etc., caused further departures from traditional timetable allocations, and in turn erased much of the rigidity associated with the mandatory use of blocks of time for each subject. In recognition and support of what were considered desirable improvements in school programs, the guidelines in 1967 to 1971 placed considerable emphasis on the need for flexibility in time-tabling.

> Every teacher is called upon to make timetables to help in meeting the needs, abilities, and interests of the pupils; no two timetables in a school need be alike; a timetable should be flexible.

The stress on greater flexibility appears to have resulted from, to have been accompanied by, and to have created less rigidity in teaching techniques and a related further increase in the use of interest groupings, individualized approaches, integrated units, and other approaches calling for special kinds of long and short-term planning.

After 1975 references to timetables shifted toward descriptions of the nature of long and short-term planning, and spelled out in greater detail references made in the 1967 to 1971 guidelines. While recognizing that much of the necessary planning must be done by the individual teacher, the new guideline stated that school staffs both as individual teachers and as a collective body under the leadership of the principal were to be the program planners. It was emphasized that "individual teachers have the responsibility of selecting strategies, resources, and activities appropriate to the needs of individual children...." In addition the cooperative effort of program planning teams was also stressed. "Certain important parts of the planning process are, however, best done cooperatively...by the entire teaching staff...by all the teachers of a given division."

(g) Pupil Progress. The guidelines issued in 1937, 1941, and 1955 for the primary and junior divisions were based on a program of studies "arranged in six successive grades or levels of attainment." Throughout this period from 1937 to 1967, and accelerating in its final years, a number of trends emerged which broke down to a considerable extent the notion of mastery in each level as the sole criterion for entry into the next level. Included in these trends were decreases in "skipping" or rapid acceleration and in non-promotion or failure; greater attention to the individual pace of learning of each student; a growing inclination to promote students on social as well as academic criteria; the introduction of alternative promotion patterns, including unit systems of promotion; the use of grouping in the classroom to accommodate different levels of attainment; the expansion of special education services; the elimination of the high school entrance examination; greater awareness of research evidence on the psychological and academic results of failure of an entire year; greater recognition of growth in creative areas as well as in academic subjects; and consideration of the overall advantages to the child of the most appropriate placement for the next year's activities. The practices resulting from these trends were recognized in documents from 1967 onward, with shifted emphasis to the continuous progress of students. Based on the assumption that "there is no set time or grade level at which every child can be expected to develop..." current guidelines note that "the wisdom of retarding a child for a year because of failure...is open to grave questions." The guideline adds that "programs should also accommodate any modifications that may be necessary to meet the needs of children with learning or other disabilities."

(h) Program of Studies: English. For the purposes of the present study two of the many changes in the guidelines regarding programs in English merit special attention. An important component of English programs, as stated in the guidelines for the period from 1937 to 1967, was supplemental reading. During the latter part of the period of the present longitudinal study, from 1967 to the present, the emphasis shifted to broader dimensions of communication, with emphasis on listening, speaking, reading, and writing. A more important change in the English program in the period from 1937 to the present has been the gradual alteration in general teaching methodologies, from an early emphasis on written exercises to the present broad variety of teaching methodologies and types of student activities and responses, including the use of other media and greater use of oral forms of expression with attendant emphasis on listening skills.

Partlow noted that in the period between 1934 and 1954 there was a change in emphasis on the meaning of words, and a shift in emphasis from vocabulary toward appreciation of language. There does not appear today to be a clear emphasis on either vocabulary studies or appreciation of language to the exclusion of the other. Elsewhere in this report evidence is presented that youngsters today use a much broader range of acceptable words in response to test items than was the case in the 1930's and the 1950's, and possible reasons for this tendency are identified.

(i) Program of Studies: Mathematics. During the period between the 1930's and the present there has been a considerable shift in major objectives and components of mathematics programs. Brown noted that mental arithmetic was reduced after the 1930's. Partlow added that there was a much more formal approach to the solution of problems in the 1930's than in the 1950's, and greater concern with applications of mathematics in the 1950's. The greatest shift in mathematics programs has had to do with increasing knowledge and changing viewpoints about the nature of mathematics. During the 1930's and well into the 1950's the emphasis was on what has been termed "corner store mathematics", with considerable emphasis on speed and accuracy in the four basic computation skills, accompanied by problems related to using simple mathematics in everyday life at home and in commerce. The rapid expansion in mathematical knowledge and its applications to a wider spectrum of human activities led to the introduction of the "new math," with less focus on basic computation skills and more concern for applications of mathematics in a wide variety of contexts, new approaches and components of mathematics, greater emphasis on problem-solving, and a view of mathematics as a model-building activity. By the late 1960's a number of new components could be identified in mathematics programs, both as emphasized in guidelines and as practised in classrooms; these included elements of geometry and set theory, emphasis on measurement, greater attention to problem-solving, the use of diagramming, and the introduction of the metric system.

The intention of the foregoing section has been to identify changes which could logically be associated with changes in student performance in the 1970's on tests designed and used in the 1930's and 1950's. The researchers are aware that numerous other changes are documented in ministry guidelines and are reflected in school practice. However, those not identified in this section were considered to have a less significant bearing on student achievement in language arts and mathematics in 1975 and 1976. Similarly, a number of current memoranda from the Ministry of Education, recent speeches of the Minister of Education, and current support documents are not commented on in this study, mainly because they post-date the termination of the current phase of the testing program.

THE EFFECTS OF LIVING AND LEARNING

In June of 1968, the Report of the Provincial Committee on the Aims and Objectives of Education in the Schools of Ontario, entitled Living and Learning, and commonly referred to as the Hall-Dennis Report, was completed and distributed to every school in the province. This report, declared to be a document for discussion rather than a blueprint for change in Ontario schools, resulted from the studies and deliberations of twenty-two citizens over a three-year period. No attempt will be made in the current study to summarize the report, or to analyze even the more important of its 258 recommendations. However, a consideration must be made in the current study of the effects that the report has had on the nature of students' school experiences in the period since 1968 and therefore on test scores in the current phase of the study.

Among those emphases and recommendations of the report which have a logical relationship to school programs and therefore are considered to have had a bearing on changes in schools and on student performance in the period since 1968 have been, among others, the following: emphasis on meeting the interests and needs of each learner; provisions for equality of opportunity; emphasis on enquiry and discovery as modes of learning; greater flexibility in time-tabling and promotion; less formal and punitive disciplinary practices; greater teacher autonomy; improvements in teacher education; the integration of curriculum into three broad areas - communications, the social sciences, and the arts; and greater attention to the pre-school and early-school years of childhood. The shift in classroom practice envisaged in the report is well summarized on page 169 of Living and Learning, as follows:

> One of the fundamental issues facing Ontario schools is the shift of focus from structured content to the child, or young person, as an individual learner. The change, already well underway, has many ramifications. The graded system, as a succession of achievement levels, will be abandoned in favor of continuous progress by the pupil - progress at his own rate in the various types of work, study, and activity appropriate for the school. The concept of passing or failing and of being promoted or made to repeat a year will disappear. What confronts the learner will not be exclusively or mainly subject matter prearranged to meet requirements of adult logic, but opportunities to pursue with zest what he

> can appreciate for its interest and value in the vibrant world of today. The schools that we envisage will give every pupil an opportunity to participate in selecting and planning his own studies.

Despite the fact that Living and Learning was not formally adopted as policy by the Ministry of Education, much of the report had an influence on subsequent guidelines and school practice. Some observers believe that the initial influence of the report was to acknowledge and give support to numerous curriculum innovations already underway. Many believe that the spirit of the report had an important impact in reducing the severity of disciplinary practices in schools. Probably the greatest impact of the report on school practice was the increased encouragement it gave to teachers to experiment with innovative programs; to explore different grouping patterns and teaching techniques; to give greater emphasis to individual students' interests and needs; to develop theme approaches and integrated units of study; and in particular to increase the amount of individual attention given to students. Its most widespread effect was the discussion on objectives and climates for learning that it motivated in almost every school community.

While no study has been done to document or evaluate the impact of Living and Learning on the education of students in the province, any reasonable historical perspective indicates that schools are different from what they were before the report was prepared, and that most of the major changes have been in the directions emphasized by the report.

The match between major emphases in the report and continuing changes in school practice are as evident in the schools of St. Catharines as elsewhere across the province. For example, while continuous progress is by no means a complete reality in these schools, it is apparent that the grade level distinctions once so highly visible no longer constitute the basis of school organization. Also the threat of year-end failure is no longer the reality that it was formerly. Moreover, the emphasis in the report on long-range planning within the comprehensive scope of a child's education and subject to the unique differences among children is now generally an intention and in many cases a reality. The report has had an effect on the insights and attitudes of teachers, not only as a result of being cited by numerous educators and professionals in journals and public speeches but also because it has been a fundamental resource for student teachers in colleges of education.

The fact that the report does not specify aims and objectives supports a trend toward making teachers and principals more aware of and more responsible for programs in their own schools, by placing the need to make decisions squarely on the shoulders of those responsible for implementing them.

Because of the impossibility of establishing the precise relationship between the report and changes in schools generally or in the specific schools in this study during the period from 1968 to 1975, it is impossible to delineate those that have occurred because of the report. However, the report remains as a fundamental document identifying current thinking about what might happen in schools. As such, it has had a profound albeit unmeasured impact on schools, even though the ultimate nature of this impact may never be clearly stated or understood. An additional footnote, suggesting that the impact of the report on education is greater than that formally recognized in public documents, is the fact that numerous educators throughout the province believe that the report is a statement of policy of the Ministry of Education, while in point of fact the report, as stated above, has never been given such formal recognition.

In attempting to account for changes in student performance today as compared to that of the 1930's and 1950's the researchers must consider, in addition to the four stated variables, a complex of interrelated influences, of which Living and Learning is an integral part.

The foregoing sections identify and discuss a large set of interrelated factors concerning changes in philosophy, objectives, supervision, methodology, and programs in schools. Together these represent one of the major variables to be taken into account in the analysis of changes in student performance; for purposes of brevity this complex of factors is collectively referred to as "changes in objectives." Two of the other major variables, socio-economic and age-grade levels, are introduced in the remainder of this chapter.

SOCIO-ECONOMIC FACTORS

A large body of research published in recent years has supported and confirmed the conclusion reached by many teachers from less formal observations, namely that there is a correlation between socio-economic conditions in the school community and student achievement. The strength of the correlation suggests a probable cause-and-effect relationship and a number of researchers go so far as to conclude that lower socio-economic conditions result in lower general levels of student performance.

Wynn (1967), in looking at the correlation between achievement and socio-economic levels, reported that

> ...(lower classes) are notoriously deficient in language skills, have had few experiences outside the home, and lack the kind of home environment which contributes to readiness for and success in school. (p. 41)

Thus, it appears that children in lower classes suffer from a learning deficit before entering school. Bernstein (1970) relates the deficit mainly to inferior language acquisition. Upon school entry, students from lower socio-economic groups must "orient towards a different structure of meanings, whether it is in the form of reading books, in the form of language use and dialect, or in the pattern of social relationships...." (p. 346)

In suggesting reasons for the correlation between lower socio-economic status and lower achievement, Napoli (1968) examined the relationship between amounts of reading material in the home and reading performance. He found that good readers had more reading material in their homes and that middle class students surpassed lower class students both in achievement and in home experiences with reading materials. (p. 555)

Froelich (1967) stated that the involvement of parents was a factor in the relationship. In speaking of a successful learning program he wrote that "the involvement of parents is an integral part of the program, so that the child feels he has support from the home." (p. 24)

Some researchers contend that lower class homes differ so markedly from school learning environments that children suffer an adjustment shock which inhibits achievement in school. Findlayson (1971) presents the relationship in this way:

> ...the children of parents with working-class values therefore do less well than the children of parents with middle-class values who foster behaviour and attitudes considered conducive to successful school performance. (p. 161)

Squibb (1973) supports the previous viewpoint, and states that, "The value system of the lower class is such that it constrains their behaviour, their self-perceptions, and their aspirations," with the result that their aptitudes and abilities become defined in such a way that "to engage successfully in middle-class behaviour is difficult." (p. 116)

Despite fairly general agreement on the correlation between socio-economic factors and student performance, researchers have been hampered by the lack of suitable means of measuring the effects of socio-economic differences. Despite this problem, in a longitudinal study such as this some treatment of the socio-economic factor becomes essential for the interpretation of comparative performance.

In order to understand the importance of the socio-economic factor, let us consider a hypothetical case in which the mean score for a grade 6 class in mathematics dropped from 35 to 30 between

1954 and 1975 and that all relative factors remained constant. From this information one could conclude that students apparently were doing worse in 1975. However, if in this example the socio-economic status of the school community dropped from middle class to lower class during the intervening years, one would conclude that the change in socio-economic level probably accounted at least in part for the lowering of mean scores. In this example, one might even attribute most of the reduction of 5 points in the average to the lowering in socio-economic status. The situation is less clear in actual situations, however, because a large set of other variables in addition to socio-economic factors is at work. The question then is to determine how one might sort out and assign the extent to which changes in socio-economic status affect changes in achievement. The approach to this problem, as employed in this study, is described in the section that follows.

In his study Dr. Partlow looked at the socio-economic changes in St. Catharines for the two decades between the first two phases of the study (1933 and 1954), and classified the schools on the basis of upper, middle, and lower class socio-economic status in 1954. The present study accepts the data in the Partlow classification, and makes a further assessment of the changes in socio-economic status of school populations since that time. Because the concern and focus of the present study is on the effects of changes that have occurred, attention is given mainly to the direction of change in socio-economic levels of the schools in question relative to the rest of the city. Observation of Table 4:1 reveals that almost all schools in this study have dropped in relative socio-economic status when compared to the populations of all schools in St. Catharines. The point must be emphasized that "below average" or "a drop in relative socio-economic status" as employed in this study do not mean a decrease in absolute terms. Rather, they mean that compared to the average rate of upward change in all school districts in the city, the change in a school described as being "below average" or having "a drop in relative socio-economic status" was less than average. Similarly "above average" refers to rates of upward change greater

TABLE 4:1

A Descriptive Classification of the Socio-Economic Status of School Populations in St. Catharines for the Period 1934-1954 and the Relative Socio-Economic Status for 1975-1976 for the Schools in the Study. School 3 no longer exists. Schools 10 and 12 did not exist in 1954.

Schools	1934 and 1954			1975 - 1976		
	Low	Average	High	Low	Average	High
1		√		√		
2	√			√		
4	√			√		
5			√		√	
6		√		√		
7		√		√		
8			√	√		
9		√		√		
10					√	
11		√		√		
12						√

than the average for the city. Thus the table shows relative changes in socio-economic status in three classes of school - those that have moved upward faster than the average, those whose change in socio-economic status was average for the city, and those which had an upward change below the average rate for the city, or which actually had a reduction or lowering in socio-economic status.

The classification of the 1975 school districts, as shown in Table 4:1 is based on information obtained from school records, income tax data, local assessments, references from law enforcement agencies, evaluations by local sociologists, teachers, principals and board of education officials, statements by agents of social and family services such as the Children's Aid Society, comparisons made by municipal government personnel including officials of the city clerk's office, and comparisons made by federal government immigration officers. Although certain minor conflicts appeared in the data collected, these differences were relatively rare, and the opinions of most of the sources of data were consistent with each other. Having sorted out the array of descriptive information, the investigator classified the schools in the study in 1975 into the three categories shown, based on the definition presented above, i.e., that such categories referred to the amount of change compared to the average change for all districts in the city. The researchers then checked this socio-economic classification once more against descriptive rankings provided by teachers and administrative staffs of the schools, board officials, municipal government personnel, and samples of parents. Once the classification system was understood, there was general agreement with the classification shown in Table 4:1.

One special characteristic of the socio-economic analysis should be identified. While the factor under study is referred to as the socio-economic status, the greater weighting in evaluations appears to have been given to economic rather than to social factors. Only a few of the persons who assisted us with these comparisons made frequent references to social conditions, while almost everyone commented on relative economic levels. This apparent imbalance in the consideration of social and economic forces creates a situation that should be recognized. Because there has been a general increase in economic conditions, purchasing power, and plane of living since 1954, it is not surprising to find that in absolute terms almost all school districts in the city were rated initially as having a higher absolute socio-economic status. This fact tended to support the logic of the investigator's plan, namely to consider changes in socio-economic factors relative to the rest of the city, rather than in absolute terms.

In relating these relative socio-economic changes to student achievement data, the investigator required a reasonably precise quantitative index of these two relationships. However, as stated earlier, while the literature is replete with correlational evidence associating lower achievement with lower socio-economic status, none of the authors provides an approach by which a valid, statistical procedure can be determined for quantitative use in predicting student achievement from socio-economic data. For this reason, the investigator has developed and introduced a novel and viable means by which, with reasonable inference, quantitative differences can be established from descriptive data about schools of differing socio-economic status.

In Tables 4:2 to 4:7 are shown the weighted means in raw scores on each of the final tests for the schools rated below average, average, and above average by the procedures described above. In order to fill in these values in some cases data were used from schools outside the study group. Observation of these tables shows in almost every case a drop in scores between above average and average, and between average and below average relative socio-economic status. From these observations one can draw reasonable inferences regarding relative effects of dropping from one position to another.

Let us now consider how these inferences would apply to an individual school described as average in 1933 and 1954 but which has dropped to below average in relative socio-economic status in 1975. By comparing the 1975 scores from this school with those from another school which was in the average group in 1975, the researcher

TABLE 4:2

Grade Means for Low, Average and High Socio-Economic Status Students of St. Catharines Public Schools for the Niagara Mathematics Test Series (Fundamental Operations), Grades V-VII (H. and M. Revision, 1976.

* Extrapolation or Interpolation

Grade	S.E.S.	Weighted Means	Number of Students
5	Low	28.93	260
	Average	29.30	80
	High	29.67*	---
6	Low	43.37	
	Average	47.57	70
	High	51.77*	---
7	Low	53.57	228
	Average	56.24*	---
	High	58.91	111

TABLE 4:3

Grade Means for Students of Low, Average, and High Scoio-Economic Status, St. Catharines Public Schools, for the Niagara Mathematics Test Series, Arithmetic Computation, Grade VIII (H. and M. Revision), 1976.

* Interpolation

Grade	S.E.S.	Weighted Means	Number of Students
8	Low	14.96	223
	Average	16.09*	---
	High	17.22	142

TABLE 4:4

Grade Means for Students of Low, Average, and High Socio-Economic Status, St. Catharines Public Schools, for the Niagara Mathematics Test Series, Arithmetic Reasoning, Grade VIII (H. and M. Revision), 1976.

* Interpolation

Grade	S.E.S.	Weighted Means	Number of Students
8	Low	7.26	221
	Average	7.55*	---
	High	7.84	144

TABLE 4:5

Grade Means for Students of Low, Average, and High Socio-Economic Status, St. Catharines Public Schools, Vocabulary Test of Common Names (Niagara Revision), 1976

* Extrapolation/Interpolation

Grade	S.E.S.	Weighted Means	Number of Students
5	Low	50.47	264
	Average	56.31	69
	High	56.79*	---
6	Low	57.61	265
	Average	59.10	66
	High	60.59*	---
7	Low	64.00	192
	Average	64.38	110
	High	71.02	135
8	Low	71.13	268
	Average	72.42*	---
	High	73.70	152

TABLE 4:6

Grade Means for Students of Low, Average, and High Socio-Economic Status, St. Catharines Public Schools, for the Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B, for 1975.

Form A

Grade	S.E.S.	Weighted Means	Number of Students
8	Low	13.06	245
	Average	13.10	475
	High	13.81	146

Form B
* Interpolation

Grade	S.E.S.	Weighted Means	Number of Students
8	Low	12.99	468
	Average	13.77*	---
	High	14.56	206

TABLE 4:7

Grade Means for Students of Low, Average and High Socio-Economic Status, St. Catharines Public Schools, for the Thorndike-McCall Reading Scale, Grade VI, 1920 Edition, Form 2, 1975.
* Extrapolation

Grade	S.E.S.	Weighted Means	Number of Students
8	Low	22.63	235
	Average	22.53	192
	High	22.63*	---

can gain a rough approximation of the effect of this drop in relative socio-economic status. In combining these approximations for all schools, the values for each school must be weighted by the size of the population writing the test. The summary of these weighted approximations for average changes in socio-economic factors for each grade in all the core schools writing the final tests is shown in Table 4:8. In considering the procedure for applying the approximations from Table 4:8 to the overall results for each test, the reader is reminded that these approximations are points in the raw score. It is of interest that the Thorndike McCall Test carries an adjustment of 0.8 points, while several tests at some grade levels require adjustments of more than one point, and for one test an adjustment of over three points is required.

The specific steps in the procedure employed in arriving at the values in Table 4:8 require further elaboration, particularly for persons wishing to employ it in other longitudinal studies. The reader is reminded that this procedure was developed in the absence of any standard procedure for quantifying socio-economic changes during longitudinal evaluations, and that caution must be exercised in the use of the procedure because it yields at best only a rough approximation of the effect of socio-economic changes. Given these qualifications, the strategy as applied to each test operates as follows:

1. Establish the socio-economic status of the schools in which classes wrote the test in the 1933 and 1954 phases of the study. (The data for 1954 are available in Table 4:1. Although Partlow referred to slight upward and downward trends from 1933 to 1954 for some of the schools, no major changes in socio-economic status for any of the schools were indicated and it is assumed that the relative socio-economic positions of the schools remained reasonably stable during this period.)

2. For each socio-economic level in 1975, determine the weighted mean raw score for each grade writing the test. These data are shown in Tables 4:2 to 4:7.

TABLE 4:8

Approximate Adjustments in Students' Raw Scores on Selected Tests Attributable to Change in Socio-Economic Status, 1954 to 1975-1976.

Test	Grade	Adjustment in Raw Score Points
Niagara Mathematics Test Series (Fundamental Operations)	5	+0.44
	6	+2.81
	7	+2.67
Niagara Mathematics Test Series (Arithmetic Computation)	8	+1.13
Niagara Mathematics Test Series (Arithmetic Reasoning)	8	+0.29
Thorndike-McCall Reading Scale	6	+0.00
Vocabulary Test (Niagara Revision)	5	+3.68
	6	+1.73
	7	+0.38
	8	+1.50
Dominion Group Achievement Test, Part 1, Paragraph Reading, Form A	8	+0.04
Dominion Group Achievement Test, Part 1, Paragraph Reading, Form B	8	+0.78

3. Calculate, by simple subtraction, the differences in weighted mean raw scores for each successive change in socio-economic status within each grade.
4. Determine from Table 4:1 the comparative status of each of the core schools in 1954 and 1975 and select the appropriate difference from the calculations in step 3.
5. Determine the number of students in each school that wrote the test.
6. Calculate the weighted value for each school by multiplying the difference selected in step 4 by the respective number of students.
7. Add the weighted values for each school (gain or loss) and divide by the total population writing the test in all core schools to determine an average adjustment score for each grade writing the test. These values for all the tests are shown in Table 4:8.

In applying the adjustment to the original scores, all that is required is the application of the adjustment to the mean raw scores for each grade on each of the final tests. This procedure is used consistently in the analysis of scores on final tests in Chapters 6 and 7.

In the application of the preceding strategy, a number of difficulties had to be overcome. For example, the investigator noticed that while several schools in the first two phases of the study were categorized as high in socio-economic status, none of the core schools in the present phase was described as above average. In order to determine a value for the movement downward from high or above average to average, additional schools outside the study group were given the tests in phase three, thereby enabling the research team to collect the data for these calculations. In the few cases where for a variety of reasons tests were not administered to schools outside the core group that were rated "high" or "average," extrapolation or interpolation was used, based on the assumption of equality of differences in socio-economic levels.

AGE-GRADE LEVELS

In any comparison of achievement levels for any one particular grade, an important variable to keep in mind is the comparative age levels of students in that grade during the various phases of the comparison. Obviously, if the average age at any grade level is significantly higher or lower during different phases of the study, this factor must be taken into account in comparing performance.

The analysis of age levels is important in longitudinal studies of this kind for another reason. When the public asks, "Do grade 6 students do as well as they did twenty years ago?", there is an implied assumption that the average age of pupils in the sixth grade has remained constant. The real intent of such a question in many instances would therefore be, "Do twelve-year-olds perform as well as they did twenty years ago, regardless of the grade level in which they are registered?"

Table 4:9 shows the average age of students in grades 5 to 8 for the schools under study on May 31st of 1935, 1954, and 1975. In comparing the figures for 1935 and 1954, the investigator must agree with the statement of Partlow, who reported that "the average ages in St. Catharines have not changed significantly during the period, particularly from 1935 to 1954." During that period the averages dropped slightly for two grades and increased slightly for the other two grades in question.

In comparing the 1954 averages with those of 1975, it is noted that there has been an increase in average age in each of the four grades under study. The reader should note that the figures given are in years, not in years and months; therefore, the figure 12.2 means 12.2 years, not 12 years, 2 months.

A consideration that one must make with respect to the age-grade averages is whether these figures are higher than one would logically expect, assuming that the average child makes normal progress, i.e., moves ahead a grade in each year, and therefore does not experience failure or skipping of grades. To obtain these figures, one

TABLE 4:9

Average Ages of Students for Grades 5, 6, 7, and 8 in the St. Catharines Public Schools of the study, May 31, 1935, 1954, and 1975.

Date	Grade 5	Grade 6	Grade 7	Grade 8
1935 (May 31)	10.9	11.6	13.0	13.6
1954 (May 31)	10.8	11.8	12.8	13.7
1975 (May 31)	11.1	12.2	13.1	14.2

look at the average age at which children enter kindergarten and then add an appropriate number of years to represent the additional grades, thereby forming a simple model of normal grade-ages.

For some years, children enrolling in kindergarten in St. Catharines public schools have been required to attain their fifth birthday by the end of the calendar year. In effect, that means that these children all are born in one calendar year, five years prior to entering kindergarten. In checking with new registrants in a number of kindergarten classes starting in September 1976, the investigator found that every child was born during the 1971 calendar year, thereby confirming that the policy gives a valid basis for the age-grade model under consideration. Assuming a reasonably even distribution of birthdates across the year, the average age of kindergarten children is approximately 5.0 years at the end of June, two months prior to their beginning kindergarten. Two months later then, on the first of September in their kindergarten year, their average age is five years two months. Nine months later, on May 31st of their kindergarten year, the pupils have an average age of five years eleven months, or 5.9 years. Based on our assumptions, then, grade 1 students on May 31st are an average of 6.9 years of age; grade 5 students, 10.9 years; grade 6 students, 11.9 years; grade 7 students, 12.9 years; and grade 8 students, 13.9 years.

Comparing these average figures with those for 1975 in Table 4:9, it is noted that the students in the schools of the study at all four grade levels are 0.2 or 0.3 years, or between two and four months, older than the averages computed in our model. The figures imply that the rate of retardation exceeds acceleration by the time students reach the grades in question, by approximately two to four months during the elementary school years. A qualification must be made with respect to these figures because they did not take into account students in special classes, wherein the average age-grade is higher than in regular classes. Obviously, then, the inclusion of these students in calculations of average age-grades would increase the figure for each year and thereby increase slightly the gap between model and actual values.

Many readers will no doubt feel that the average age at the end of the eighth grade is unexpectedly high, and will point out quite correctly that it is considerably higher than the "model" and even higher when compared to the figures for 1935 and 1954. The stated policy in many school systems, including the one under study, and the general rhetoric about promotion in schools, indicate that children are promoted when their achievement, ability, and social development indicate they are ready to be moved ahead. Sometimes the general policy does not refer to promotion per se, but instead states that the program that each child experiences is in keeping with his achievement, abilities, and social development. The latter is merely a paraphrasing of the earlier statement, in view of the fact that an advancement in program ultimately implies an advancement or acceleration in grade level. A common assumption derived from these stated policies is that youngsters, on the average, are promoted in a pattern matching or slightly ahead of their chronological growth, and that therefore there would probably be a slight overall increase in acceleration over retardation. Obviously this is not the case in the study schools, at least as demonstrated by average age-level data, and therefore the assumptions concerning promotion by achievement or "continuous progress" are inconsistent with actual practice.

In comparing the 1975 figures for the four grades in question with those in the model described above, one notes that the policy of retarding or accelerating youngsters in keeping with their abilities seems either not to apply or to have little effect on the averages at the grade 5 to 8 level, in view of the fact that comparisons from grade to grade show almost exactly a one-year average increase from grade to grade. Therefore, by and large the slight average retardation in these schools takes place almost entirely in the first five years in school.

A comparison of the "model" averages for grades 5 to 8 with the data in Table 4:9 shows not only that the students in the study schools today are on average older than the "model" values, but also that students in 1935 and 1954 were younger than these values. Further explanation is needed to account for these differences.

Studies of the age-grade levels for primary grades stated in the Brown and Partlow reports indicate that the average age in grade 1 was as low as 6.2 years, or about 0.7 years younger than the figures in the model described above. This figure alone would more than account for the difference shown in Table 4:9. An apparent reason for the difference between the 6.2 and the 6.7 years in the early studies and in the model, respectively, has to do with the availability of kindergarten in all schools. Prior to the introduction of kindergarten, the successful interventions of parents to gain approval for admitting seemingly precocious children to school at an earlier age had the effect of lowering average ages at the grade one level. Another obvious explanation is in the difference in acceleration practices in the different periods in question. As stated earlier, it was a common practice, particularly in the 1930's, to have as many as thirty to forty percent of the students take the second Book (grades 3 and 4) in one year. Obviously this would lower the average age in grades 5 to 8. An examination of Table 4:9 will show that even at the grade 5 to 8 level the rate of progress differed sharply in the three periods under study. In 1935, pupils moved from grades 5 to 8 in a total of 2.7 years, or 0.9 years per grade. In 1954 they progressed through these grades in a total of 2.9 years, or an average of 0.97 years per grade, while in 1975 they moved from grade 5 to grade 8 in a total of 3.1 years, or an average 1.03 years per grade. One of the factors accounting for the difference in these figures has been the reduction in "skipping" and in some other forms of acceleration in favour of enrichment of the program within the existing grade structures. Many teachers today are convinced that lateral enrichment of the program is preferable to acceleration, and the application of this viewpoint in the classes under study may account for the fairly uniform match between average age increases and grade levels.

For a wide variety of reasons, related both to provisions in school programs and to the lack of opportunities for employment outside of school, fewer of the older and less able students now leave school at ages 13 to 16. Early school leaving is not tolerated in school systems today to the extent that it was a generation or two ago. Obviously the retention in grade 7 and 8 classrooms of a greater number of older students has a marked upward effect on the average age.

Some of the increase in age-grade levels can also be attributed to socio-economic factors. The reader will recall from the socio-economic analysis above that this factor shows lower levels, relative to the rest of the city, in the schools under study. On the basis of socio-economic factors alone, then, one would anticipate a slight increase in average age-grade levels for the schools under study. In Lady Churchill School, a senior public school in a newer area in the southern part of the city, the average age in grade 8 is 13.8, as compared to 14.2 for the study schools. One could reasonably expect that the average age-grades for the entire city would be somewhat below the 1975 figures in the core schools.

The researchers in the present study used an extremely thorough checking procedure to determine the ages of all pupils accurately. In doing so, numerous inconsistencies were found between student ages reported on the Ontario Student Record cards and on the test summary sheets. In almost every instance in which an error occurred, the student summary sheets indicated that students were younger than the ages reported on the O.S.R. cards. The investigator hypothesizes that these errors occur mainly for two reasons - difficulty on the part of some students in converting age into years and months, and a deliberate attempt, particularly by students older than the average age for the grade, to make themselves appear younger than they in fact were. Although the research cannot assess the original data in these respects, an inference can be made that similar clerical and procedural errors probably also existed in the earlier phases of the study, and that the existence of these "honest errors" would tend to lower the reported average age-grade slightly.

Whatever the reasons for the increase in average ages in all the grades under study, the fact remains that these differences do exist. The changes in this variable must be taken into account in any analysis of comparative student performance.

The fact that several of the tests were written by students at all four grade levels presented the investigator with a simple and logical procedure for obtaining an approximation of the changes in raw scores attributable to differences in average ages. From the mean test scores for each grade the gains from grade to grade were determined; because of the uniformity of age-grade and grade progression it can be concluded that the grade progressions match increases in age. Therefore once the gain for a grade is determined, the gain for a year is also established. From this value the amount of gain attributable to the slight increase (0.3 for grades 5 and 7, 0.4 for grade 6, and 0.5 for grade 8) in age-grade from 1954 to 1975 or 1976 for each such test can be calculated.

For those tests written by only one grade a modified and slightly less defensible procedure was employed to gain an approximation of the effect of age-grade changes. In these instances the age difference as a proportion of the actual mean age was applied to the mean score to establish the approximation of the effect of age difference. The values obtained in this way were compared with those arrived at by the earlier formula in order to ensure that the approximations appeared to be roughly comparable. Table 4:10 shows the age-grade differences from 1954 to 1976, with the number of points in the mean scores on each final test attributable to age-grade differences alone, obtained by using the procedures described above.

In the case of the Thorndike-McCall Reading scale, written by grade 6 students only, a procedure was employed that was more defensible than the latter adapted procedure. The availability of a comprehensive data base for this test enabled the investigator to determine normative values for each age level, and therefore to calculate expected gains or losses for differences in average grade-ages.

TABLE 4:10

Approximate Adjustments in Students' Raw Scores on Selected Tests Attributable to Changes in Mean Age-Grade, 1954 to 1975-76.

Test	Grade	Change in Mean Age-Grade 1954 to 1975-76	Number of Points (Raw Score) Attributable To Changes in Mean Age-Grade
Niagara Mathematics Test Series (Fundamental Operations) Grades V-VIII, 1976	5	0.3 year	3.6
	6	0.4 year	4.8
	7	0.3 year	3.6
Niagara Mathematics Test Series, Arithmetic Computation, Grade VIII, 1976	8	0.5 year	0.53
Niagara Mathematics Test Series, Arithmetic Reasoning, Grade VIII, 1976	8	0.5 year	0.25
Thorndike-McCall Reading Scale, 1920 Edition, Form 2	6	0.4 year	0.75
Vocabulary Test Number 1, Ninety Common Names (Niagara Revision), 1976	5	0.3 year	2.1
	6	0.4 year	2.8
	7	0.3 year	2.1
	8	0.5 year	3.5
Dominion Group Achievement Test, Part 1, Paragraph Reading, Grade VIII, 1934 Edition, Form B.	8	0.5 year	0.46

5 Background to the Data Analysis

Before beginning this longitudinal comparison and analysis of student performance in mathematics and language arts in the 1930's, the 1950's, and the mid-1970's, the author examined a variety of issues relating to the two earlier studies by Dr. Brown and Dr. Partlow and considered their relationship to the present study.

Common knowledge of empirical procedures suggests that the rationale for beginning a research study affects the decisions made throughout the project. For this reason, in a project of this kind, where one is conducting parallel studies for purposes of simple replication of conditions, comparison of results, and longitudinal tracking of subjects, it becomes essential that the new phase of the study identify and where possible incorporate the fundamental intentions of the original investigators. These goals or intentions, the subtle nuances that emerge from them, and the practices employed to reflect them must all be taken into account if the most recent phase is to have an objective and comparative relationship to its earlier counterparts. The nature of the research questions, the methodology employed, the alteration and administration of instruments, the statistical analysis used, and the recommendations and conclusions are inextricably related to the rationales of the respective authors. Therefore, the researcher finds it useful in this report to set forth certain considerations outlined in the sections that follow.

(a) Rationale for the Study by Dr. C.A.Brown (1930's).

According to Dr. Brown, in discussions in May 1976, the original study in the 1930's was planned and implemented with the primary purpose of introducing testing to the schools. Before the study was conducted, the St. Catharines public school system, like other school districts in the province of Ontario at that time, had never experienced a massive battery of selected, standardized tests. Therefore while the results of the testing would be of interest to Dr. Brown, in view of his administrative position in the school system, for the information they would give him in comparing achievement of his students with normative values, his primary purpose was to train his teachers in and introduce the system to the use of standardized testing procedures. Knowledge of Dr. Brown's experience and his qualifications at the time suggests that it was logical for him to have introduced a program of teacher-training in test administration, and also suggests that such a program could not have been conducted by a more competent agent in that decade. Also, while the results of the tests were obviously of interest and value at the time to Dr. Brown for comparative purposes, it is equally clear that he did not have the primary intention of conducting the tests so that future comparative studies could be made. Both Dr. Brown and Dr. Partlow indicate that the availability of the data from the 1930's was a fortuitous event in the 1950's, rather than the result of an intention to provide a data bank for future comparison.

(b) Rationale for the Study by Dr. H.R. Partlow (1950's).
The relative weighting of the two intentions in Dr. Brown's study, the introduction of a testing program and the study of achievement with data from normative tests, was reversed in the case of Dr. Partlow's phase of the study. While no doubt Dr. Partlow was also interested in training and encouraging teachers in the use of tests, including normative tests, as evidenced by the thoroughness of the memoranda for test administration which he used in the conduct of his study, it is evident that his major intention was to obtain data for use both in comparisons with the norms and, more important, in comparisons with achievement in the 1930's. As with Dr. Brown in the earlier phase, Partlow was one of the best qualified researchers in any school system to carry out this intention. Additionally, with the data from the earlier study available to him he was in a unique position to make such a comparative study.

During the period immediately prior to the Partlow study of the mid-1950's, there had been considerable debate in local and provincial newspapers and magazines concerning how well the schools were performing. Two contradictory and perhaps extreme quotations, from Maclean's magazine (March 1, 1952), and from the Globe and Mail, Toronto (April 8, 1952) respectively, respectively, exemplify the terms of the debate.

> Modern methods of teaching and discipline do not demand the best from students either in work or in behaviour. The school seems to have acknowledged its inability to cope with the harmful but powerful influence of the moving picture, the vulgar phonograph record, the comic strip, and the tawdry degrading magazine and book. Schools no longer invite their student to hitch their wagon to a star (Maclean's magazine).

The preceding quotation stands in sharp contrast to the more complimentary response from the Toronto Globe and Mail:

> There is little evidence to indicate that schools are not doing a better job now than they did a generation ago. In fact there is much evidence to suggest that they are doing a better job. Today's children read more quickly and with a greater understanding. In many parts of the city schools are havens of refuge and help for both children and adults. They assist in setting a pattern of life and raising the tone of the community (Globe and Mail).

Neither the standards of students' performance nor the effects of changes in the educational system during the 1940's and early 1950's had been evaluated in a formal, comprehensive manner prior to the study conducted by Dr. Partlow. According to Partlow, in 1954, "There have been many changes in Ontario's educational system during the last twenty years, including changes in teaching methods and textbooks, but of all these changes the introduction of the new program in 1937 and 1938 by the central authority has probably had the most far-reaching results." Dr. Partlow's study, then, appears to have been motivated primarily by two interrelated considerations concerning comparative student performance. First, he wished to gather data in order to assess the comments made in the press at that time. Secondly, he wished to assess the effects of changes in education that had occurred during the previous two decades. Unlike the Brown study, the data and analysis were not only for purposes of the researcher and the school system, but also for a wider audience; as reported earlier, the study was presented in thesis form to the University of Toronto and was published as a book.

As in the current study, Partlow identified a wide range of variables in the culture which may have had an effect on student achievement in the 1950's. He provided in some detail descriptions of the four variables dealt with in the present study - changes in school objectives, socio-economic factors, vocabulary, and age-grade levels. However, in assessing the factors bearing directly on student performance he placed by far the greatest weight on and attention to changes in school programs as the major factor contributing to changes in student performance.

As with the Brown study, Partlow did not have a deliberate intention of providing a data bank for future comparative studies. This is one of the respects in which the present study differs from both the Brown and Partlow phases. The current study recognizes and incorporates the desirability of providing a substantial bank of raw and tabulated data, and of storing it for use by future investigators. Obviously, this intention has influenced many aspects of the data collection, tabulation, and analysis.

In order to compensate for the lack of raw data from the original phases of the study, the present investigator discussed these phases thoroughly with the two earlier researchers, and as a result of these conversations and extrapolations from the earlier tables of statistical data he was able to make a realistic match between the assessments of learning in the 1930's and 1950's and those in the present phase.

(c) Instrumentation in the 1930's and 1950's. When Dr. Brown made his selection of tests to measure learning in the 1930's he did not have at his disposal the rich array of standardized instruments available to educators and researchers today. However, from the modest set of available instruments he selected and used those tests that appeared a reasonable match to the Canadian idiom, the school curriculum, and the research questions in which he was interested. In view of the fact that his major purpose in the study was to introduce teachers to the use of standardized tests, the clarity of instructions for administration obviously was also taken into account. The selection of the tests was accompanied by a certain amount of tentativeness on the part of the researcher, with respect to the quality and validity of the instruments. In May 1976 Dr. Brown recalled that he had believed in 1937 that the vocabulary test was "not too useful," it was questionable even at that time. In his report Partlow confirmed that Dr. Brown had questioned the use of the vocabulary test in 1937. Although Dr. Brown made some modification of the vocabulary test in order that it would reflect vocabulary of the time, there was no other modification of the tests in order that they should match curricula in the schools. While some consideration of the suitability of the tests was given, the Brown study tells us essentially how well students performed on the tests rather than on stated school objectives of the time.

The follow-up study by Dr. Partlow, on the other hand, clearly established the extent to which items on the tests used by Dr. Brown in the 1930's matched curriculum content for both the 1930's and the 1950's. As a result of the item analyses made by Partlow, the data presented by Brown and Partlow, and first-hand statements of these two researchers, we can resonably conclude that the statistics recorded in the 1950's represented accurate appraisals of student performance not only on the tests themselves but in achievement of school objectives of the time.

In Chapter 3 of this report the researcher presents not only the original tests but also descriptions of the procedures used in matching these tests against modern school objectives and the procedures by which modern counterparts of these tests were prepared which would match stated school objectives today.

(d) Familiarity of Staff and Students with Standardized Testing. References have already been made in this report to questions of student readiness, teacher attitudes, and training in the use of standardized tests. The differances in these respects between the respective phases of the study are of interest and probably have some bearing on performance scores.

According to Brown, testing in the 1930's was "a new game." Most of the teachers had not previously administered standardized tests. Students were familiar with classroom tests, usually constructed, administered, and marked by the classroom teacher. Brown reported that while the use of standardized tests by the teachers was a new experience, there was very little if any criticism of the testing procedure in the 1930's. One can reasonably conclude, then, that students' tests scores would not be inflated in the 1930's as a result of skill or practice with standardized tests, a factor sometimes referred to in the literature as "test wiseness" (sic). The one exception to this general assessment of practice effects would be at the Grade 8 (Senior Fourth) level. The extreme importance and attention given to the High School Entrance Examinations caused teachers to prepare their students thoroughly for such external tests. Various types of practice were employed, including teaching items from previous years' examinations, inserting such items in class tests, analyzing question types for frequency of use, and even simulations in which students wrote entire sets of previous examinations. Published sets of these examinations were in use in most Senior Fourth classrooms.

In the 1950's students and staff members were familiar with standardized tests. The use of such tests had become a familiar if not common practice. Routine tests of subject matter were regularly conducted by teachers. Partlow also reported that principals and teachers showed a complete willingness to participate in his study, an indication that the general procedures were familiar to them. Negative attitudes or comments, if they existed, were not documented.

Since the 1950's there has been an increasing unfamiliarity with standardized tests and a growing reaction against them. The elimination of the high school entrance examination prior to 1950 and of provincially administered secondary school graduation examinations a decade later reduced the general emphasis on external examinations. During the 1960's and 1970's books such as the Tyranny of Testing reflected and affected teachers' growing dislike of testing procedures which rendered them only limited descriptions of the learner. More recently a number of studies of intelligence quotients, including the well known studies by Jansen, and findings related to socio-economic class differences and race differences in intelligence quotients, have posed serious challenges to standardized testing. Even so, by the time an elementary school student has reached the 8th grade he or she has normally experienced a number of tests, standardized or constructed by the teacher. Many of these are administered to evaluate individual differences, the effectiveness of teaching practice, proper placement of students, etc., in addition to students' performance. However, it is very unlikely that students in the grades tested would have had sufficient practice with standardized instruments similar to the ones used in the study to introduce any significant element of familiarity that would inflate their "actual" scores. To control for the possibility of differences in teachers' familiarity with students and tests, the current testing was administered largely by the research team of Niagara Centre.

(e) Statistical Sophistication. As the researchers would have expected, the nature of the statistical analysis in the three phases of the study differs both because of the rationales employed and because of the progressive history of statistical analysis.

In his study in the 1930's Dr. Brown analyzed the data in a manner that would enable him to achieve the two goals of his study - training the teachers in administering and using standardized tests, and interpreting the data to provide him with both a reasonably comprehensive picture of learning standards in the classrooms of the third and fourth Books in the public schools of St. Catharines and an analysis of the factors that accounted for the changes identified.

Dr. Partlow's intention in the 1950's to add a comparative analysis to the study brought about a change in the statistical procedures. However, again it appears that relatively little thought was given to the possibility of a follow-up study in the 1970's. Therefore typical statistical procedures of the time were used, which would tender the comparative information required by the researcher and enable him to make comparisons between the 1930's and 1950's

The intention of the current study not only to make a comparative analysis of changes in student performance and to attribute them to certain important variables, but also to provide a supply of data for future studies, has led the investigator to use certain statistical procedures and other practices that were not characteristic of the earlier two studies. Some of these procedures and practices are:

1. the original raw data are retained, along with all tabulations and analyses;
2. in general means are used as measures of central tendency, whereas medians were used almost exclusively in the earlier studies;
3. measures of dispersion are recorded; and
4. greater emphasis is placed on the distinction between statistical and practical significance.

In attempting to meet the most important intention of the current study, namely to compare student performance in the 1930's, 1950's, and 1970's, it is imperative that as much as possible of the data from the earlier studies, along with the statistical procedures used, be captured and employed in the present analysis. Therefore, in making the present analysis, three criteria are employed in comparisons with the earlier studies:

1. maximum use is made of the descriptive statistics from the earlier studies;

2. where possible, means replace medians as measures of central tendency for the original data; and
3. data from the 1970's are recorded for comparative purposes in a format similar to the original data.

Also, to enable future researchers to use the increasingly more comprehensive and powerful statistical procedures available to them, a complete description of the raw data will be retained.

6 Comparative Achievement in Mathematics

The purpose of this chapter is to present and analyze the data on comparative student achievement in mathematics in the three phases of the study. For purposes of convenience and clarity, the first section deals exclusively with achievement on the grade 8 tests, and the second half of the chapter on the tests written by grades 5, 6, and 7.

It will be recalled that the author's intention was to make a fair comparison of arithmetic computation and reasoning across the three phases of this longitudinal study. In the field of mathematics, all the tests employed in the 1934 phase and the 1954 phase were employed again in the 1975 phase. In addition, based on information provided in Chapter 3, and further referred to in this chapter, it was determined that replacement tests were required to reflect modern objectives in mathematics. The results of these tests are shown in additional tables for comparative purposes. It was concluded that changes in vocabulary did not have a measurable effect on differences in student performance in mathematics between the three phases of the study, and therefore no special adjustment for differences in vocabulary was required.

The other two factors, socio-economic effects and age-grade levels, however, were sufficiently important to require adjustments to the scores on the replacement tests. These adjustments, based on the formulae described in Table 4:8 for socio-economic factors and Table 4:10 for age-grade factors, increase the 1975-76 scores for socio-economic effects and depress them for age-grade level effects.

GRADE 8 ACHIEVEMENT IN ARITHMETIC COMPUTATION AND REASONING

It will be recalled, from Table 3.3, that the grade 8 students in the 1975-76 period were asked to write six mathematics tests. The names of these tests, and the dates when they were written, are presented again in Table 6:1, this time in the sequence in which they will be treated in this chapter.

Table 6:2 presents the comparison, with gains and losses in the mean raw scores for all grade 8 classes in arithmetic computation and reasoning as measured by the two Dominion Group Achievement Tests, 1934 Edition, Form A. The table shows that the highest levels of performance on both tests were in the 1952 phase, and that the 1975 raw scores are below the 1934 Ontario norms and are the lowest on both tests for any of the time periods under study. Considering, then, the raw scores alone on these two tests, it is apparent that the 1952 achievement was ahead of the 1934 norms, but that in the period from 1952 to 1975 there was a drop in performance to a point lower than the 1934 norms for these tests. The percentage of correct answers rose from a mean of 64.4% on the Ontario 1934 norms to 70.0% in 1952 and then fell off to 46.8% in the 1975 phase.

Table 6:3 presents the same picture for the two tests of the same components using Form B, 1934 Edition. The similarity of the achievement scores is not surprising, when one considers that Forms A and B are parallel forms of the same general test. As might be expected, the results in Table 6:3 are remarkably similar to those in Table 6:2. Once again we see that the 1953 raw scores were higher than the Ontario 1934 norms, and that the 1975 scores were lower than either of these. Again the percentages of correct answers rose from the normative value of 64.4% in 1934 to 68.9% in 1953, and dropped to 38.13% in 1975.

From the above analysis it is evident that grade 8 students in the 1975 phase of the study performed much worse on the 1934 tests on these two components than did their counterparts in 1934 and 1952-53.

The evidence that the grade 8 students in 1975 did worse on the 1934 tests than their counterparts in 1934 and 1952-53 should not of itself lead one to conclude that students today perform worse in mathematics than did their counterparts a generation or two ago. Before such a general conclusion would be justified one must examine the nature of the tests and

TABLE 6:1

Tests Written by Grade 8 Students in the Three Phases of the Study.

*Refers to Ontario Urban Norms.

Name of Test	Testing Dates		
	1933-38	1952-54	1975-76
Dominion Group Achievement Test, Part II, Arithmetic Computation, Grade VIII, 1934 Edition, Form A.	June 1934*	June 1952	June 1975
Dominion Group Achievement Test, Part II, Arithmetic Reasoning, Grade VIII, 1934 Edition, Form A.	June 1934*	June 1952	June 1975
Dominion Group Achievement Test, Part II, Arithmetic Computation, Grade VIII, 1934 Edition, Form B.	June 1934*	June 1953	June 1975
Dominion Group Achievement Test, Part II, Arithmetic Reasoning, Grade VIII, 1934 Edition, Form B.	June 1934*	June 1953	June 1975
Niagara Mathematics Test Series, Arithmetic Computation, Grade VIII (H. and M. Revision), 1976.	----	----	June 1976
Niagara Mathematics Test Series, Arithmetic Reasoning, Grade VIII (H. and M. Revision), 1976.	----	----	June 1976

TABLE 6:2

A Comparison of Achievement Scores in Arithmetic Computation and Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1952, 1975) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form A.

FORM A	Ontario Norms 1934	1952		1975		Gain or Loss 1934-1952	Gain or Loss 1952-1975	Gain or Loss 1934-1975
		Number of Pupils	Mean Raw Score	Number of Pupils	Mean Raw Score			
1. Arithmetic Computation (Max: 25)	18.0	347	19.2	318	15.59	+1.2	-3.61	-2.41
2. Arithmetic Reasoning (Max: 20)	11.00	347	12.3	372	5.15	+1.3	-7.15	-5.85
Total Arithmetic (Max: 45)	29.0	694	31.5	690	20.74	+2.5	-10.76	-8.26
Percentages	64.4%		70.0%		46.08%	+5.6%	-23.92%	-18.32%

TABLE 6:3

A Comparison of Achievement Scores in Arithmetic Computation and Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1953, 1975) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form B.

FORM B	Ontario Norms 1934	1953		1975		Gain or Loss	Gain or Loss	Gain or Loss
		Number of Pupils	Mean Raw Score	Number of Pupils	Mean Raw Score	1934-1953	1953-1975	1934-1975
1. Arithmetic Computation (Max: 25)	18.0	343	19.6	258	12.74	+1.6	-6.86	-5.26
2. Arithmetic Reasoning (Max: 20)	11.0	343	11.4	307	4.42	+0.4	-6.98	-6.58
Total Arithmetic (Max: 45)	29.0	686	31.0	565	17.16	+2.0	-13.84	-11.84
Percentages	64.4%		68.9%		38.13%	+4.5%	-30.77%	-26.27%

determine the degree to which they reflect objectives of the mathematics program in the three phases in question. In other words, the question of "fair" tests comes under consideration.

This issue, of whether the objectives of the mathematics program remain constant from one phase to another, was first addressed by Partlow in his study. The major changes in curricula between 1934 and 1952, according to Partlow, were in the direction of greater flexibility for the purpose of teaching according to the needs of the pupils. There was greater emphasis on diagnosis of individual difficulties and on remedial teaching to overcome these difficulties. However, these changes, in the opinion of Partlow, did not make the various forms of the Dominion Tests of Arithmetic Computation and Reasoning inappropriate for comparisons between the two periods. Partlow demonstrated that these tests gave a fair comparison of grade 8 achievement in computation and reasoning by matching the components of the respective arithmetic courses for 1934 and 1952.

The present investigator, in his introduction to the Niagara Test Series in Chapter 3, established that the 1934 editions of the Dominion Group Achievement Tests of Arithmetic Computation and Reasoning are not well matched to the curricula of today. To make valid, reliable comparisons of grade 8 student achievement in arithmetic computation and reasoning across all phases of the longitudinal study it was necessary to construct replacement tests. The two new tests, The Niagara Mathematics Test Series, Grade 8 Computation and Grade 8 Reasoning, can be examined in the appendices of this report. These two tests were administered in 1976 after the investigator had tested the suitability of the Dominion Tests for grade 8 curricula in 1975.

Table 6:4 compares the 1976 results in the two Niagara Mathematics Test Series for Computation and Reasoning with the results from Form A of the Dominion Group Achievement Test in 1934, 1952, and 1975. Table 6:5 repeats this process, making its comparisons with Form B of the Dominion Group Achievement Tests. Inspection of Tables 6:4 and 6:5 shows that students in 1976,

TABLE 6:4

A Comparison of Achievement Scores in Arithmetic Computation and Arithmetic Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines - 1952, 1975, 1976) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form A, and the Niagara Mathematics Test Series.

TEST	1934 Dominion Test	1952 Dominion Test		1975 Dominion Test		1976 Niagara Math. Test	
	Ontario Norms	No. of Pupils	Mean Raw Score	No. of Pupils	Mean Raw Score	No. of Pupils	Mean Raw Score
1. Arithmetic Computation (Max: 25)	18.0	347	19.2	318	15.59	478	15.10
2. Arithmetic Reasoning (Max: 20)	11.0	347	12.3	372	5.15	478	7.06
3. Total Arithmetic (Max: 45)	29.0	694	31.5	690	20.74	956	22.16
Percentages	64.4%		70.0%		46.08%		49.24%

TABLE 6.5

A Comparison of Achievement Scores in Arithmetic Computation and Arithmetic Reasoning for Grade 8 Students (Ontario - 1934, St. Catharines, - 1953, 1975, 1976) as Measured by the Dominion Group Achievement Tests, 1934 Edition, Form B, and the Niagara Mathematics Test Series.

TEST	1934 Dominion Test	1953 Domionion Test		1975 Dominion Test		1976 Niagara Math. Test	
	Ontario Norms	No. of Pupils	Mean Raw Score	No. of Pupils	Mean Raw Score	No. of Pupils	Mean Raw Score
1. Arithmetic Computation (Max: 25)	18.0	343	19.6	258	12.74	478	15.10
2. Arithmetic Reasoning (Max: 20)	11.0	343	11.4	307	4.42	478	7.06
3. Total Arithmetic (Max: 45)	29.0	686	31.0	565	17.16	956	22.16
Percentages	64.4%		68.9%		38.13%		49.24%

writing the Niagara Mathematics Test Series which was established as being a fair comparison with the Dominion Tests in the earlier two phases of the study, achieved less than their counterparts in 1934 and 1952 in both arithmetic computation and reasoning. The 1976 scores in the Niagara Mathematics Test Series are only slightly higher than for the 1975 results on the Dominion Group Achievement Tests. The percentage of correct answers for the Ontario norms in 1934 was 64.4%; for the 1952-53 groups it was 68.9% and 70.0% on the two forms; and in the Niagara Test Series in 1976 the percentage of correct answers was 49.24%.

The use of the comparison test establishes that the consistent drop in student performance between 1953 and 1975-76 cannot be attributed to changes in objectives alone. After adjusting for the other two major variables under study, socio-economic factors and age-grade levels, the reason for the drop in performance is still unexplained. Table 6:6 shows the adjustments in average raw scores in the 1976 administration of the Niagara Mathematics Test Series for socio-economic effects and for the change in age-grade levels. These two factors have largely a counterbalancing effect on each other, with the result that the mean raw scores for 1976 are still consistently below the 1934 and 1952-53 values.

Dr. Partlow was able to show that the Dominion Group Achievement Test was a suitable test of what grade 8 students studied in 1934 and 1952-53. The present investigator established that the same test did not adequately assess the breadth of the arithmetic program in 1976. The replacement test used in 1976 included items related to current course content in arithmetic, but obviously, by emphasizing only computation and reasoning, neglected other important dimensions of mathematics programs today. Thus the comparisons in Tables 6:2 to 6:6 do not establish that students today perform worse than the students of 1934 and 1952-53 in all aspects of arithmetic achievement, but they do establish fairly clearly that the students tested do achieve much less in the two areas of mathematics under consideration The results do not reveal that the student of 1975-76 may be

TABLE 6:6

A Comparison of Achievement in Arithmetic Computation and Reasoning for St. Catharines Grade 8 Students Based on the Dominion Group Achievement Test (Part II, Arithmetic Computation and Reasoning, Grade VIII, 1934 Edition Forms A and B) for 1934 and 1952-53 and the Niagara Mathematics Test Series (Arithmetic Computation and Reasoning, Grade VIII; H. and M. Revision, 1976) for 1976.

Achievement Area	1934 Ontario Norms		1952-53 Means (Raw Score)		1976 Unadjusted Means	1976 Adjustment For Socio-Economic Factor	1976 Adjustment For Age-Grade Factor	1976 Adjusted Means
	Form A	Form B	Form A	Form B				
Arithmetic Computation (Maximum Raw Score = 25)	18.0	18.0	19.2	19.6	15.10	+1.13	-0.53	15.70
Arithmetic Reasoning (Maximum Raw Score = 20)	11.0	11.0	12.3	11.4	7.06	+0.29	-0.25	7.10

learning aspects of arithmetic not even taught during the previous phases of the study and not assessed by the original tests or the test of "fair" comparison.

Focusing on the drop in arithmetic computation and reasoning, it is necessary to consider other factors which may account for this reduction in student achievement. A number of these additional factors are treated at considerable length in Chapter 8 and later in this chapter.

Among others, the most apparent factor has been the reduction in the overall proportion of time and emphasis given to mathematics in the total curriculum, brought about by increased emphasis on a wide range of other subjects, and a resulting reduction in the proportionate emphasis on mathematics. Another major factor appears to be the result of a shift in methodology away from considerable use of drill, oral computation, numerous short tests, and other practices which tended to produce high levels of computation skill and accuracy. The influence of high school entrance examinations reinforced such methodology in the 1930's and had residual influences in the 1950's.

The inappropriateness today of numerous items in both the computation and reasoning components of the 1934 tests, including aspects of mathematics no longer considered important or useful, accounts for a reduction in student scores in the Dominion Group Achievement Tests. However, it is more difficult to explain why students did not perform at appropriate significantly higher levels when these aspects of mathematics were replaced by more appropriate ones in the Niagara Test Series.

An anomaly that cannot be explained on the basis of the foregoing conclusions alone has to do with the comparative performance in computation when compared to results in reasoning. The explanations presented above imply that students today should have comparatively better success in reasoning tests than in computation tests, when compared to their counterparts in 1934 and 1952-53. However, Tables 6:4 and 6:6 show that this was not the case, and that in both the 1934 Dominion Tests and the 1976 Niagara Tests the students performed worse in 1975-76 in reasoning than in computation, when compared to their counterparts in the earlier two phases. Reasons for this anomaly are not apparent. However there is a possibility that less attention is paid today than formerly to established procedures for solving problems containing several steps, of the kinds exemplified in the tests. Obviously this was a highly developed skill in the 1930's, particularly at the grade 8 level, because of the presence of the grade 8 examinations which featured problem-solving questions of this kind. No doubt much of this influence remained well into the 1950's. In any event Partlow did not comment on any reduction in the formal approach to solving mathematical problems of this type.

FUNDAMENTAL ARITHMETIC OPERATIONS, GRADES 5 to 7

It will be recalled from Chapter 3 that the students in grades 5 to 7 in all three phases of the study wrote the Dominion Arithmetic Test of Fundamental Operations, Grades 4 to 8, 1934 Edition, Form A Also, the classes in these grades in the core schools in 1976 wrote the Niagara Mathematics Test Series (Fundamental Operations), Grades 5 to 7, 1976. The scores on these two tests form the basis for the analysis that follows. Because these two tests were considered of particularly high quality, and covered a wide range of arithmetic objectives, and because their application over three consecutive grade levels added a vertical dimension to the comparisons, considerable attention is given to the comparative results.

The rationale for developing and using the Niagara Mathematics Test Series (Fundamental Operations), Grades 5 to 7, 1976 was introduced in Chapter 3 and will be developed at greater length later in this chapter. First, however, a comparison of performance on the original test, the Dominion Arithmetic Test of Fundamental Operations, Grades 4 to 8, 1934 Edition, Form A is justified.

Tables 6:7, 6:8 and 6:9 show class and city medians for

TABLE 6:7

Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), for the Dominion Arithmetic Test of Fundamental Operations, 1934 Edition, Form A, for Grade Five Classes of St. Catharines Public Schools, 1936, 1954, 1975.

Grade 5 Classes

1936 N = 412 (Class Median)	1954 N = 452 (Class Median)	1975 N = 298 (Class Mean)
21	30.5	18.06
21	29	17.93
20	25.5	17.92
20	23.5	17.51
20	23	16.77
19	22	16.02
19	22	15.93
19	21.5	14.48
19	21	
19	21	
18	21	
17	19	
15	19	
14	18	
	18	
City Median 1936 = 19	City Median 1954 = 22	Weighted City Mean 1975 = 16.75

TABLE 6:8

Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), St. Catharines Public Schools, Dominion Arithmetic Test of Fundamental Operations, Grade 6, 1934 Edition, Form A, 1936, 1954, 1975.

Grade 6 Classes

1936 N = 448 (Class Median)	1954 N = 438 (Class Median)	1975 N = 269 (Class Mean)
30	41.5	28.92
29	41.5	27.19
29	41	25.27
27	41	24.67
26	39	24.06
26	38	23.64
26	36	21.92
26	36	21.81
25	33.5	
25	33	
25	33	
23	32	
23	30	
22	30	
City Median 1936 = 26	City Median 1954 = 36	Weighted City Mean 1975 = 24.46

TABLE 6:9

Ranked Class Medians and City Medians (1936, 1954), Class Means and a Weighted City Mean (1975), St. Catharines Public Schools, Dominion Arithmetic Test of Fundamental Operations, Grade 7, 1934 Edition, Form A, 1936, 1954, 1975.

Grade 7 Classes

1936 N = 417 (Class Median)	1954 N = 410 (Class Median)	1975 N = 496 (Class Mean)
50	57.5	39.58
50	56	36.57
50	55	31.24
49	53	30.40
47	51	29.5
46	49	29.48
45	46.5	25.96
44	46	25.03
44	46	
37	45	
37	43	
34	42	
	39.5	
City Median 1936 = 45	City Median 1954 = 49	Weighted City Mean 1975 = 32.53

1936 and 1954 and class means and weighted city means for 1975 for grades 5, 6, and 7 respectively. The rationale for using means in 1975-76 is presented in Chapter 5. The data in these tables are presented in this form in order to retain as closely as possible certain aspects of the test results for each class, thereby providing for the reader evidence of the distribution from class to class and school to school across the study group.

An inspection of these tables indicates clearly that the class means in 1975 were considerably below the medians in 1936 and even lower than the medians in 1954. In all three grades the highest class means in 1975 were similar to the lowest class medians in 1954. In all three grades the weighted means of the entire core group in 1975 were below the city medians in both 1936 and 1954, although in two cases they approached the 1936 values. From the summaries of class scores presented in these three tables there can be no doubt that the students in the core schools performed worse on this test than did their counterparts in the earlier two phases of the study.

A more finely grained analysis of the achievement scores for these three grades across the three phases of the study reveals that the lower pattern in 1975 is not consistent across all dimensions of mathematics, and thereby raises serious questions about the comparability of the tests with respect to objectives and the weighting and balance among various components. Tables 6:10, 6:11, and 6:12 show the percentages of questions missed in each phase of the study in the eleven content groups in the test. In the first five groups of questions, which included the four basic forms of computation and weights and measures, the 1975 students did consistently worse than their counterparts in 1936 and 1954. In the four fundamental operations with fractions the grade 5 pupils in 1975 performed better than their counterparts in 1936, but less well than those in 1954. Grade 6 and 7 students in 1975, however, performed worse than both earlier groups in these operations. In the four operations with decimals the 1975 grade 5 students outperformed

TABLE 6:10

Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 5 Classes: St. Catharines, 1936, 1954, 1975.

Type of Question	No. of Questions	% Missed 1936	% Missed 1954	% Missed 1975	Gain or Loss 1936-1954	Gain or Loss 1954-1975	Gain or Loss 1936-1975
Addition	5	8.0	7.9	12.55	+0.1	-4.65	-4.55
Subtraction	5	12.2	12.7	27.25	-0.5	-14.55	-15.05
Multiplication	5	18.5	24.1	45.37	-5.6	-21.27	-26.87
Division	5	29.1	33.5	56.44	-4.4	-22.94	-27.34
Weights and Measures	10	81.4	75.6	87.65	+5.8	-12.05	-6.25
Fractions: Addition and Subtraction	10	99.6	76.6	89.56	+23.0	-12.96	+10.04
Fractions: Multiplication and Division	10	96.7	90.5	93.56	+6.2	-3.06	+3.14
Decimals: Addition and Subtraction	5	97.5	93.5	90.47	+4.0	+3.03	+7.03
Decimals: Multiplication and Division	5	99.8	98.3	92.82	+1.5	+5.48	+6.98
Percentage	10	99.9	99.3	99.43	+0.6	-0.13	+0.47
Percentage and Decimals	10	100.0	99.9	99.76	+0.1	+0.14	+0.24
TOTALS	80	76.3%	72.1%	79.05%	+4.2%	-6.96%	-2.78%

TABLE 6:11

Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 6 Classes: St. Catharines, 1936, 1954, 1975.

Type of Question	No. of Questions	% Missed 1936	% Missed 1954	% Missed 1975	Gain or Loss 1936-1954	Gain or Loss 1954-1975	Gain or Loss 1936-1975
Addition	5	7.6	5.3	11.45	+2.3	-6.15	-3.85
Subtraction	5	10.1	9.4	19.25	+0.7	-9.85	-9.15
Multiplication	5	12.9	14.0	34.12	-1.1	-20.12	-21.22
Division	5	22.1	19.7	41.78	+2.4	-22.08	-19.68
Weights and Measures	10	62.6	58.0	80.26	+4.6	-22.26	-17.66
Fractions: Addition and Subtraction	10	64.1	46.9	77.21	+17.2	-30.31	-13.11
Fractions: Multiplication and Division	10	86.0	67.4	80.30	+18.6	-12.9	+5.7
Decimals: Addition and Subtraction	5	93.8	27.5	57.47	+66.3	-29.97	+36.33
Decimals: Multiplication and Division	5	97.9	61.3	73.75	+36.6	-12.45	+24.15
Percentage	10	99.1	95.9	95.65	+3.2	+0.25	+3.45
Percentage and Decimals	10	99.9	99.1	99.33	+0.8	-0.23	+0.57
TOTALS	80	70.02%	54.49%	68.96%	+15.53%	-14.47%	-2.22%

TABLE 6:12

Percentage of Questions Missed and Relative Gains or Losses for Each Type, Dominion Arithmetic Test of Fundamental Operations, Grade 7 Classes: St. Catharines, 1936, 1954, 1975.

Type of Question	No. of Questions	% Missed 1936	% Missed 1954	% Missed 1975	Gain or Loss 1936-1954	Gain or Loss 1954-1975	Gain or Loss 1936-1975
Addition	5	3.9	6.2	10.12	-2.3	-3.92	-6.22
Subtraction	5	7.4	8.6	15.20	-1.2	-6.6	-7.8
Multiplication	5	10.5	13.4	28.51	-2.9	-15.11	-18.01
Division	5	16.9	17.6	37.18	-0.7	-19.58	-20.28
Weights and Measures	10	48.5	39.8	68.43	+8.7	-28.63	-19.93
Fractions: Addition and Subtraction	10	43.4	32.8	59.37	+10.6	-26.57	-15.97
Fractions: Multiplication and Division	10	39.3	34.4	72.90	+4.9	-38.5	-33.6
Decimals: Addition and Subtraction	5	36.0	13.9	38.39	+22.1	-24.49	-2.39
Decimals: Multiplication and Division	5	51.4	30.7	49.23	+20.7	-18.53	+2.17
Percentage	10	75.9	66.1	85.24	+9.8	-19.14	-9.34
Percentage and Decimals	10	96.1	92.8	97.84	+3.3	-5.04	-1.74
TOTALS	80	45.8%	38.9%	59.14%	+6.9%	-20.25%	-13.35%

their counterparts in both earlier phases, and grade 6 students obtained higher scores than their 1936 counterparts. In the two groups of questions involving percentages, the 1975 students were generally better than the 1936 students, and outperformed the 1954 students in two cases.

Table 6:13 combines the percentages of each content grouping missed in all three grades, and presents the overall percentages for comparative purposes. Inspection of this table shows that the slightly higher performances of grades 5 and 6 in 1975 on certain groupings are erased in the total picture to a considerable degree, suggesting that students today may learn some of these components of mathematics at a slightly earlier age level, but that their proficiency does not continue to surpass that of the grades in the earlier phases of the study once they had been exposed to these components. The table shows that the students in 1975 missed a greater percentage of each type of question than did the students in 1954, and also missed overall a greater percentage than students in 1936 in all the groupings except operations with decimals.

The apparent anomalies in the data in the foregoing tables raise questions concerning the suitability of the 1934 test for students in 1975, in that they imply changes in the grade levels at which certain content groupings are introduced. This concern points to the need for consideration of "fair" comparisons, a fundamental aspect of the design of the current study.

Before administering the Dominion Test in 1975 the investigator studied its suitability in the earlier two phases of the study. Both Dr. Brown and Dr. Partlow administered this test, after having judged its suitability. In deciding in 1954 to use the same test as that employed by Dr. Brown in 1936, Dr. Partlow established that the nature of arithmetic curricula for the grades in question was not significantly changed in the intervening years. He did report, however, that there were some differences in recommended time allocations in 1954 as compared to 1936. The suggested time allotment in 1936 was 165 minutes,

TABLE 6:13

Percentage of Each Type of Question Missed: Dominion Arithmetic Test of Fundamental Operations (80 Questions), All Classes, Grades 5, 6, 7: 1936, 1954, 1975.

		Percentage of Questions Missed		
		1936	1954	1975
Group 1	Addition (5)	6.5	6.5	11.13
Group 2	Subtraction (5)	9.9	10.3	19.60
Group 3	Multiplication (5)	13.9	17.3	34.65
Group 4	Division (5)	22.6	23.8	43.74
Group 5	Weights and Measures (10)	64.0	58.4	76.81
Group 6	Fractions: Addition and Subtraction (10)	68.7	52.8	72.34
Group 7	Fractions: Multiplication and Division (10)	74.2	65.0	80.56
Group 8	Decimals: Addition and Subtraction (5)	76.1	46.2	57.81
Group 9	Decimals: Muitiplication and Division (5)	83.3	64.5	67.65
Group 10	Percentage (10)	91.7	87.7	91.85
Group 11	Percentage and Decimals (10)	98.7	97.4	98.75
TOTAL	80 Questions	62.9	55.7	67.2

175 minutes, and 175 minutes per week for grades 5, 6, and 7 respectively. The comparable time allotments for the same grades in 1954 were 150 minutes for grade 5; 150 minutes for grade 6; and 210 minutes for grade 7. Recognizing that there were only modest time differences in the two periods and that actual practice in schools varied from the suggested amounts, Partlow assumed that there was approximately the same amount of time spent on fundamental operations in 1936 and 1954. Further, he accepted the Dominion Test as a fair measure of student achievement in arithmetic in both 1936 and 1954.

For reasons identified in Chapter 3, the present investigator found that the descriptions of the relative similarity in arithmetic content and therefore the suitability of the Dominion Test as a "fair" instrument in 1975 did not apply. The introduction of the "New Mathematics" in the period since 1954, large increases in the emphasis on understanding rather than on overlearning of skills, and the addition to the mathematics program of numerous components that did not exist in arithmetic curricula in the two earlier time periods all represent major changes in objectives in program in 1975 when compared to the earlier two periods. Further, a shift in the basic nature or purpose of mathematics, from what has been called a "corner-store" approach in the earlier two periods to a concept of mathematics as a system for developing models to interpret quantitative data in many activities of modern life, represents a major change in focus as well as content. These observations and judgments about lack of suitability of the 1934 tests for assessing performance in today's mathematics curriculum were supported in interviews with classroom teachers, mathematics specialists, and consultants in mathematics for school boards and the Ministry of Education. The investigator contends therefore that in order to establish a "fair" comparison between performance in 1936, 1954, and 1975, the comparisons should be based on results from the Dominion Test for the earlier two phases and from a comparable test suited to modern mathematics programs, such as the Niagara Mathematics Test, in the current phase.

Tables 6:14, 6:15, and 6:16 show the class medians on the Dominion Tests for 1936 and 1954, and the class means on the Niagara Test in 1976. Instead of making comparisons only from the data in these three tables, the reader should also look at the data in Table 6:17, where the means in 1976 are adjusted for the other two major variables under study, namely the socio-economic factors and the age-grade factors. It will be noted that the socio-economic factors provide an upward adjustment for the core schools at all three grade levels, and that the age-grade factor applies a reduction at all three grade levels, resulting in a slight overall reduction from the unadjusted means in all three cases.

Comparing either unadjusted means in 1976 or the adjusted means for the same year with the medians in 1954 and 1936, one finds an overall increase in scores between 1936 and 1954, and again between 1954 and 1976.

Another feature of the 1976 class means is worthy of note. Tables 6:14 to 6:16 show that the range of class means in 1976 is much greater than the range of medians in either 1936 or in 1954. The top-ranked class medians in 1976 are much higher than their counterparts in 1936 and 1954, but the lowest medians in 1976 were only slightly better, and in one case lower, than the lowest median in the earlier two time periods. These data support a number of other observations made in this study, including statements about the greater heterogeneity of modern classes, the disinclination to accelerate brighter students, the retention in classes of slower students as a result of social promotion, and effects of different relative socio-economic conditions. Other factors may also be at work in creating the very wide range in class means, and officials at the local level who may wish to consider the implications of this study may also wish to determine the causes for such a wide discrepancy in average performance at identical grade levels in different schools.

From the data in Tables 6:14 to 6:17, the investigator

TABLE 6:14

Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A, 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematics Test Series (Fundamental Operations).

Grade 5

1936 N = 412 (Class Median)	1954 N = 452 (Class Median)	1976 N = 296 (Class Mean)
21	30.5	40.92
21	29	37.43
20	25.5	35.56
20	23.5	35.3
20	23	32.96
19	22	29.8
19	22	28.72
19	21.5	28.11
19	21	27.64
19	21	27.07
18	21	25.5
17	19	24.0
15	19	23.43
14	18	23.36
	18	17.36
City Median 1936 = 19	City Median 1954 = 22	Weighted City Mean 1976 = 29.72

TABLE 6:15

Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A; 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematics Test Series (Fundamental Operations).

Grade 6

1936 N = 448 (Class Median)	1954 N = 438 (Class Median)	1976 N = 315 (Class Mean)
30	41.5	52.96
29	41.5	46.26
29	41	45.25
27	41	43.81
26	39	42.9
26	38	40.85
26	36	39.52
26	36	38.22
25	33.5	37.07
25	33	34.93
25	33	34.04
23	32	34
23	30	32.75
22	30	31.44
City Median 1936 = 26	City Median 1954 = 36	Weighted City Mean 1976 = 40.46

TABLE 6:16

Ranked Class Medians and City Medians (1936, 1954), Dominion Arithmetic Test of Fundamental Operations, Grades 4-8, 1934 Edition, Form A; 1936 and 1954; Ranked Class Means and Weighted City Mean (1976), Niagara Mathematics Test Series (Fundamental Operations).

Grade 7

1936 N = 417 (Class Median)	1954 N = 410 (Class Median)	1976 N = 451 (Class Mean)
50	57.5	67.7
50	56	67.3
50	55	66.2
49	53	63.3
47	51	56.5
46	49	55.0
45	46.5	53.4
44	46	51.3
44	46	50.9
37	45	48.0
37	43	47.9
34	42	47.7
	39.5	47.4
		47.2
		46.0
		45.0
City Median 1936 = 45	City Median 1954 = 49	Weighted City Mean 1976 = 53.57

TABLE 6:17

A Comparison of Achievement in Arithmetic Fundamental Operations for St. Catharines Grades 5, 6, 7 Students Based on the Dominion Arithmetic Test of Fundamental Operations (Grades IV to VIII, 1934 Edition, Form A) for 1936, 1954 and the Niagara Mathematics Test Series (Fundamental Operations), Grades V-VIII (H. and M. Revision), for 1976.

Grade	1936 (Medians)	1954 (Medians)	1976 (Weighted Means: Unadjusted)	1976 Adjustment For Socio-Economic Factor	1976 Adjustment For Age-Grade Factor	1976 Adjusted Means
	Form A	Form A				
5	19	22	29.72	+0.44	-3.6	26.56
6	26	36	40.46	+2.81	-4.8	38.47
7	45	49	53.57	+2.67	-3.6	52.64

concludes that when students in 1976 in these schools were tested with an instrument which could be described as a "fair" test, that is a test which reflected the objectives and emphases in modern mathematics programs, the results when compared with student performance at the grade 5 to 7 levels in 1936 and 1954 were approximately the same, and apparently slightly higher.

In comparing the results for grades 5 to 7 with the results reported earlier for grade 8, the investigator is left with a dilemma. How is it that grades 5, 6, and 7 apparently do better in 1976 in fundamental operations than their counterparts a generation or two ago, while grade 8 students do consistently worse in 1976 than their earlier counterparts in arithmetic computation and reasoning? A number of plausible explanations to this paradox are quickly apparent. First, while the match between the Niagara Test of Fundamental Operations and curricula for grades 5, 6, and 7 may be adequate, such is not the case for the grade 8 arithmetic curriculum of 1976. Both Dominion Group Achievement tests and the Niagara Test Series in Computation and Reasoning deal only with these two aspects of the grade 8 program today. While these two dimensions of mathematics represented the major part of the mathematics curriculum in the two earlier phases, they receive less overall emphasis today, and proportionately less time and attention.

A second possibility is that in an attempt to select items which are more characteristic of actual practice in grade 5 to 7 classes today, the designers selected items which in one way or another were actually easier than their counterparts in the earlier tests. For example, the earlier tests used a number of items in such areas as multiplication and division which employed numbers with far more digits than are normally employed in classroom exercises and tests today. To the extent that the mere size of the numbers contributed to the difficulty of correct computation, the modern test may be judged to be slightly easier.

A third explanation underlines further the difference in emphasis on grade 8 mathematics between the former periods and today. In 1976, the emphasis on mathematics when compared to other subjects appears to be roughly the same as in the other grades in the upper half of the elementary school. On the other hand, in the earlier phases of the study, and particularly prior to the removal of the high school entrance examination at the end of the eighth grade or its earlier equivalent (senior fourth), and continuing well into the 1950's as a result of some tardiness in altering practice after the removal of these external examinations, there was an extreme emphasis on mathematics achievement in grade 8, to an extent out of all proportion to that in the earlier grades. Among the extreme practices employed in some schools to ensure success in grade 8, particularly in mathematics, some of which were referred to earlier in this report, were the addition of an extra hour to the school day, usually prior to 9 a.m.; the formal teaching and testing of previous questions or entire examinations; the practice of classifying previous types of questions and predicting the types likely to appear on the next set; formal instruction in how to apply stereotyped solutions to two- and three-step problems in areas such as interest, taxation, discount, etc.; the use of extended periods of written or oral drill aimed at overlearning of number facts and computation skills; the use of homework to increase practice; the retention in grade 7 of certain students who were not expected to pass the entrance examinations; and a number of other equally effective practices. With the gradual removal of these forms of additional emphasis on reasoning and computation at the eighth grade level, it could be expected that levels of performance at this grade level would return to more or less the levels common to grades 5, 6, and 7. However, the Dominion Tests of Arithmetic Reasoning and Computation developed in 1934 reflected achievement levels resulting from the emphasis described above, and therefore provided normative values far in advance of what would be expected when such extreme emphasis was removed at the grade 8 level.

A fourth explanation has to do with the introduction of the rotary system at the seventh and eighth grade levels. The rotary system, now in common use, controls the amount of time that can be allocated to mathematics or any other subject. Prior to the introduction of this system, individual teachers could decide on time allocations to different subjects, and under the pressures referred to above, naturally gave greater time as one of the forms of increased emphasis on grade 8 mathematics.

In spite of these four plausible explanations, the investigator is not completely convinced that grade 8 scores in arithmetic and computation are as high today as one might expect. On the other hand, the students in grades 5 to 7 in the study obtained higher scores than did their counterparts in the earlier two phases, using a test of similar construction and reasonably comparable difficulty, but employing questions related to modern content and objectives of mathematics programs. The mean scores for these grades, when adjusted slightly downward as a result of the introduction of factors accounting for socio-economic changes and changes in age-grade level, remained slightly higher than those in the earlier two phases of the study. The researcher, while not convinced that grade 8 scores in the two mathematics areas tested are at achievement levels that might be expected, concludes that parents, trustees, and the public in general have relatively little reason for serious concern with achievement levels in mathematics in the schools under study. Having considered such factors as changes in school objectives, socio-economic conditions, and age-grade levels, the investigator concludes that students at the grade 8 level perform somewhat worse in arithmetic computation and reasoning as measured by a test of fair comparison than their counterparts in the earlier two phases of the study; however, students in grades 5 to 7 perform at least as well and probably somewhat better on comparable "fair" tests in mathematics than did students in these grades a generation or two ago.

ITEM ANALYSIS: THE NIAGARA MATHEMATICS TEST SERIES, REVISED EDITION, FUNDAMENTAL OPERATIONS, 1976.

For a number of reasons, the results from grades 5, 6, and 7 on the Niagara Mathematics Test Series, Revised Edition, Fundamental Operations, 1976, merit further analysis, both with respect to the insights they provide about student performance on the items in the test, and as a means of evaluating the test itself as a valid instrument for use today. The investigator has already presented reasons for the construction of this test, and a general description of provisions to ensure its rational validity. An analysis of student performance across the grades on the test will further document the validity of the test items and will illustrate the desirable quality of being able to assess growth across the grades in each test item.

Figures 6:1(a) to 6:1(d) present in graphic form the mean scores of students in the three grades on each of the eighty items of the test. A careful study of those graphs shows that, with the exception of items 59 and 80, and four easy questions at the beginning of the test, every item shows growth across the three grades. The exceptional items all are at the extremes of student ability, and therefore were influenced either by general mastery or by a very small number of correct responses, some of which resulted from guessing. In only three questions, items 9, 16, and 73, was there an overlap of mean scores, the result of high general mastery across the grades in two cases, and of successful guessing by a few students in the other.

The expected trend, in which higher grades achieve higher scores on each item, demonstrates the rational consistency of the content of the test. The graph also indicates the items introduced at various grade levels, because in general such new items are identified with sizeable gains in scores. The items showing the greatest intergrade spread between grades 5 and 6 are 30, 60, 71, and 72. Those showing a large spread between grades 6 and 7 are 20, 25, 30, 32, 34, 71, and 72.

The graph also indicates that there is a reasonably consistent increase in item difficulty as students proceed through the grades. For example all grade averages are high for items 1 to 14. For items 14 to 80, with the exception of items 59, 73, and 80, there is a general consistency in the degree of apparent difficulty. Thus for example the entries for item 14 show that the percentages

Figure 6:1(a)

The Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976.
Percentages of correct responses for each item at each grade level.

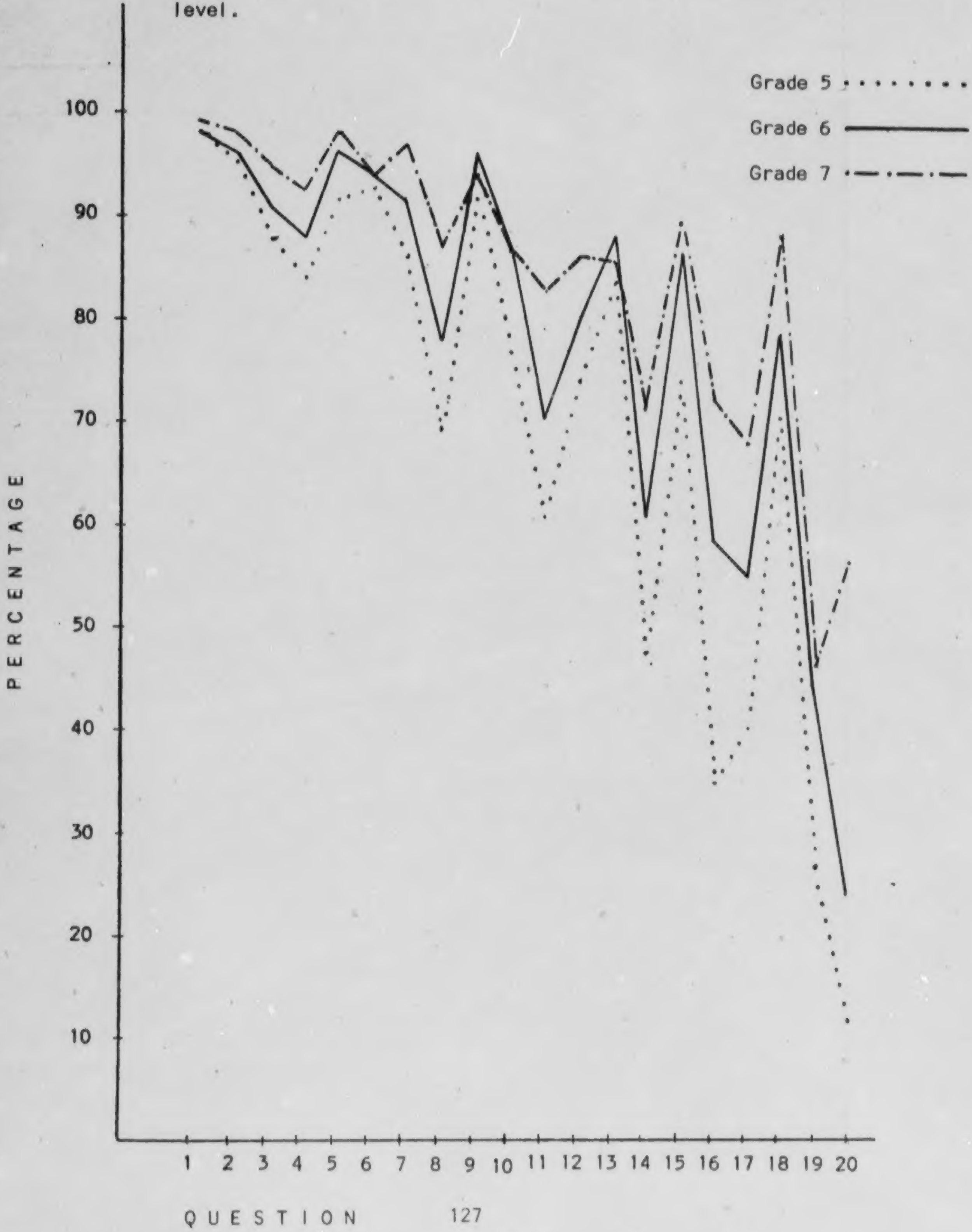


Figure 6:1(b)

The Niagara Mathematics Test Series (Fundamental Operations) Grades V to VII (H. and M. Revision), 1976.

Percentages of correct responses for each item at each grade level.

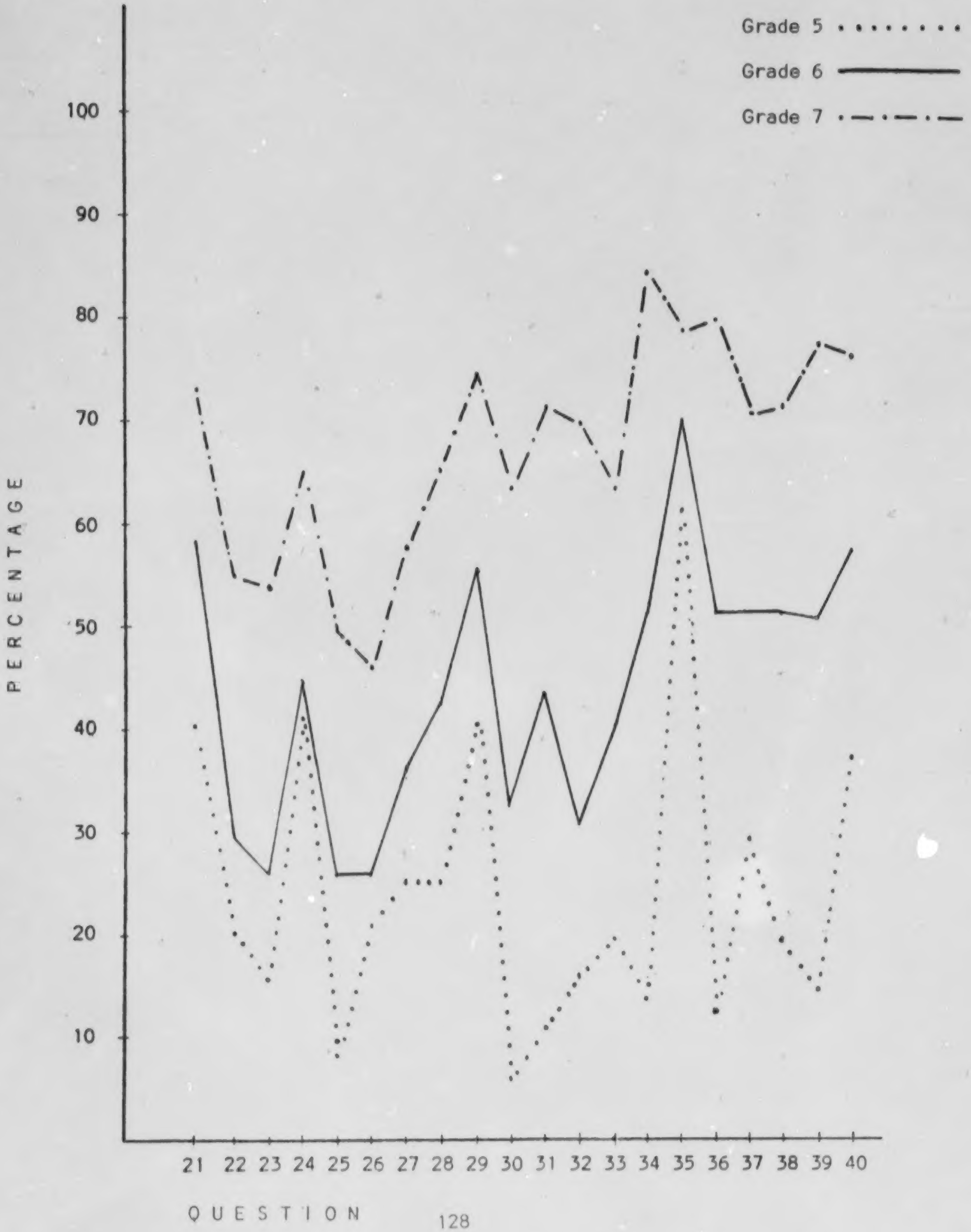


Figure 6:1(c)

The Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976.
Percentages of correct responses for each item at each grade level.

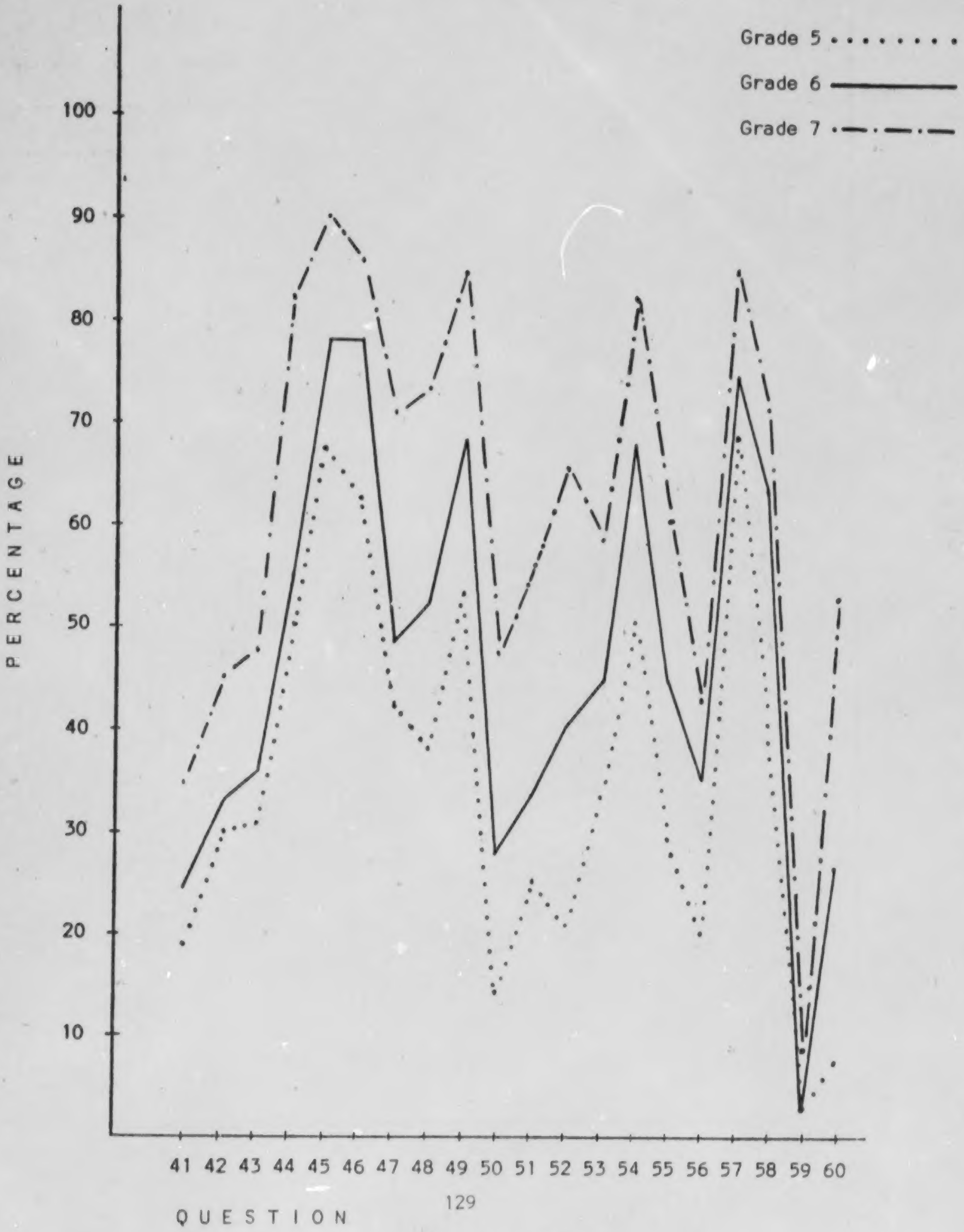
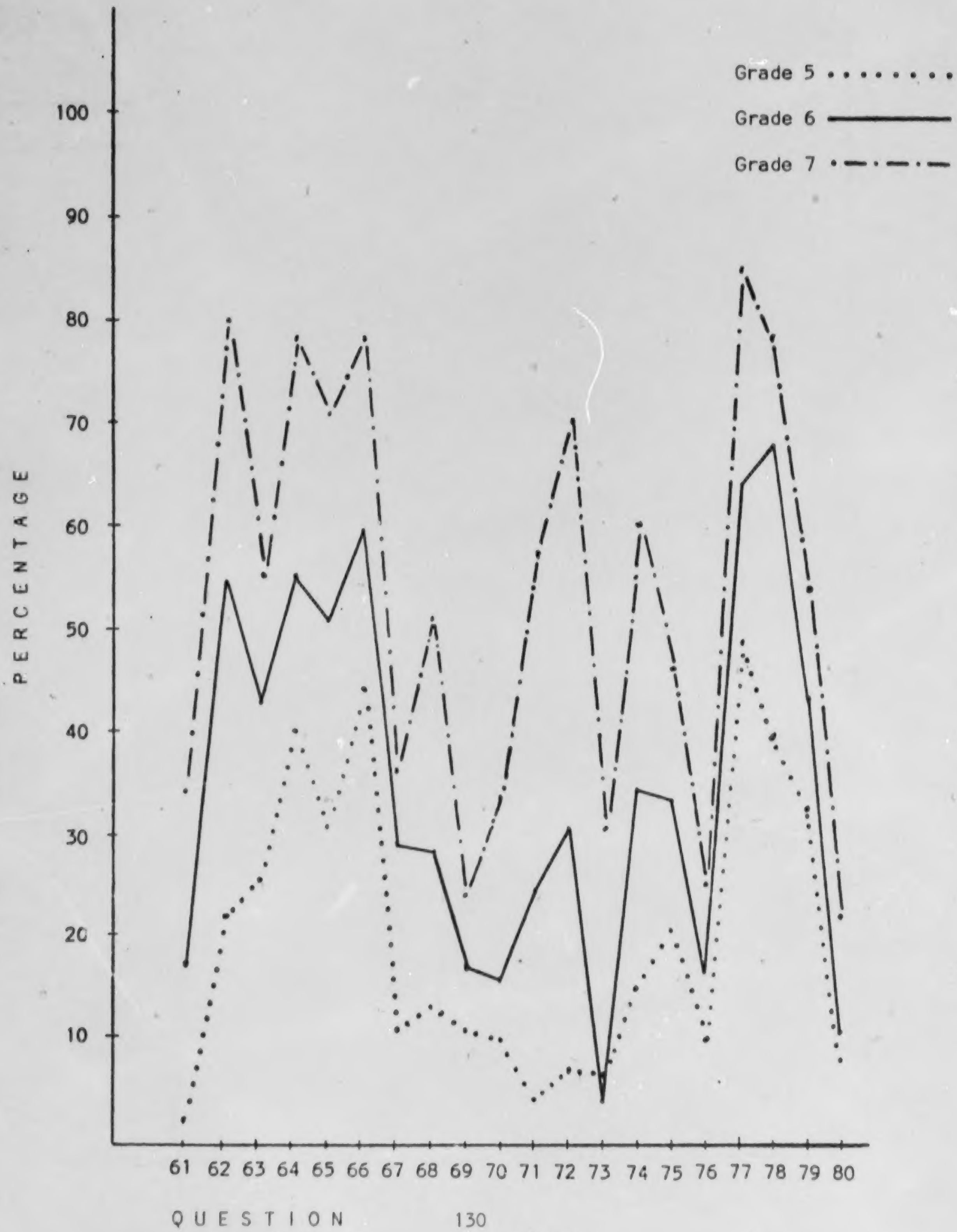


Figure 6:1(d)

The Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976.

Percentages of correct responses for each item at each grade level.



of students answering this item correctly in grades 5, 6, and 7 respectively are 47, 61, and 71. Item 70, near the end of the test shows percentages of 10, 16, and 33 for the respective grades.

The graph assists the investigator in another respect, by identifying items which appear to have resulted in atypical performance for all three grades, a situation which makes the items themselves suspect. Four such items are numbers 59, 61, 73, and 80. By referring to this test in the appendices, the reader may wish to evaluate the possible reasons presented below for the apparent mismatch or unsuitability of these items.

Item 59 appears at first to be a simple fraction and whole number operation. However, the question breaks stride in the test, as it were, from previous questions in that the student is not told in words to subtract. With no directing instructions, the student was apparently left to determine alone what was required in the operation. The low scores on this item indicate confusion rather than inability to subtract. At lower grade levels it seems likely that many students also rejected the question because it required a negative answer. This intention of the question, however, seems to have been obliterated by the apparent confusion it caused over the operation to be employed.

Item 61 appears to be a simple item requiring multiplication of a fraction and a mixed number. However, only 2 percent of grade 5 students, 16 percent of grade 6 students, and 34 percent of grade 7 students answered this question correctly, suggesting that it was confusing rather than difficult. The fact that it did not specify that the student was to multiply accounts for some of the confusion. The similarity in the size of the type face for the whole number and the fraction apparently confused other students.

Item 73 represents an extremely simple example of the use of lowest common multiples. The low scores appear to indicate that students in grades 5 and 6 do not understand the term "common multiple," but obviously a reasonable number of grade 7 pupils have either learned the term or were able to figure out what was required from the context of the question.

Item 80, seemingly a straightforward item, clearly was confusing to many students. In fact only 22 percent of students at the grade 7 level answered this question correctly. Further analysis of students' responses indicated that answer A was selected most frequently. Apparently many students took the question at face value, believing the statement implied that Mr. Brown was always going to take groups of five in his car. Some students saw this as a trick question and focused their thinking on the supposed hidden trick rather than on the fairly obvious correct solution. The expectation of difficult concluding items as a result of experience in writing similar tests caused some students to select inappropriate answers, perhaps against their better first judgment. Obviously the item requires revision since students appear to have been penalized by the wording of the question.

Tables 6:18(a) to 6:18(d) and Tables 6:19(a) to 6:19(d) provide further data for the analysis of the test. The first set of tables shows the number of correct responses in each grade for each question and the percentages of incorrect responses. The second set of tables presents the percentage, for each grade, of errors in eight groups of items. The questions are grouped into appropriate categories, and then the percentages of questions missed by each grade on each question and on the group are presented. The tables readily show the aspects of mathematics in which students have low, medium, and high levels of error, and conversely, those in which levels of mastery are high, medium, and low. Of greater value are the insights the tables provide into the reductions in error in each group from grade to grade and the apparent effects of introducing and teaching the skills in question. By analyzing these data and examining the test items in each group (see appendices) the reader can determine whether the students perform to the levels one might expect. The researcher feels obligated to point out the inconsistency across the four primary operations, and particularly the high percentages of error in multiplication and division. Other areas of concern are problem-solving, all operations with fractions, and the combination of decimals with fractions and percentages. While the amount of error in some of these categories at the grade 5 level could be expected, these percentages should drop more rapidly after the skills involved appear in the curriculum in grades 6 and 7, and presumably are taught to the classes involved.

Further analysis and revision will be required before this test could be considered "standardized." Also, split-half coefficients could readily be found in order to determine the extent to which superior students respond differently from less able students. Considering the overall quality of the test, as reflected by the validity of the items and the pattern of intergrade differences as shown in Figures 6:1(a) to 6:1(d) and Tables 6:19(a) to 6:19(d), the author postulates that the Niagara Mathematics Test Series, Fundamental Operations, 1976, is a reasonable modern facsimile of the Dominion Test it replaced, and with further improvements would be highly useful in future as a test of fundamental operations in grades 5 to 7.

TABLE 6:18(a)

Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976. Number of Correct Responses and Percentage of Incorrect Responses for each Question. Grade 5: n = 296; Grade 6: n = 315; Grade 7: n = 451.

Question No.	Grade 5		Grade 6		Grade 7	
	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect
1	289	2	309	2	446	1
2	280	5	302	4	440	2
3	260	12	286	9	428	5
4	250	16	276	12	418	7
5	271	8	303	4	443	2
6	274	7	296	6	422	6
7	256	14	289	8	436	3
8	205	31	246	22	392	13
9	271	8	303	4	422	6
10	228	23	273	13	392	13
11	180	39	219	30	376	17
12	218	26	251	20	387	14
13	249	16	278	12	385	15
14	138	53	191	39	321	29
15	218	26	272	14	403	11
16	105	65	182	42	326	28
17	118	60	169	46	304	33
18	208	30	245	22	397	12
19	79	73	137	56	209	54
20	36	88	76	76	251	44

TABLE 6:18(b)

Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976. Number of Correct Responses and Percentage of Incorrect Responses for **each Question.** Grade 5: n = 296; Grade 6: n = 315; Grade 7: n = 451.

Question No.	Grade 5		Grade 6		Grade 7	
	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect
21	118	60	182	42	330	27
22	58	80	90	71	250	45
23	48	84	82	74	245	46
24	120	59	138	56	295	35
25	24	92	83	74	266	51
26	61	79	83	74	207	54
27	74	75	112	64	258	43
28	73	75	133	58	294	35
29	117	60	174	45	334	26
30	18	94	103	67	283	37
31	32	89	135	57	320	29
32	48	84	99	69	313	31
33	55	81	126	60	286	37
34	38	87	161	49	380	16
35	182	39	217	31	353	22
36	36	88	162	49	358	21
37	87	71	162	49	317	30
38	57	81	161	49	318	29
39	42	86	158	50	348	23
40	110	63	179	43	345	24

TABLE 6:18(c)

Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976. Number of Correct Responses and Percentage of Incorrect Responses for each Question. Grade 5: n = 296; Grade 6: n = 315; Grade 7: n = 451.

Question No.	Grade 5		Grade 6		Grade 7	
	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect
41	57	81	75	76	154	66
42	89	70	104	67	203	55
43	93	69	114	64	212	53
44	149	50	174	45	370	18
45	198	33	246	22	408	10
46	187	37	247	22	390	14
47	125	58	150	52	321	29
48	111	62	165	48	330	27
49	158	47	214	32	380	16
50	40	86	85	72	210	53
51	74	75	105	67	254	44
52	61	79	126	60	291	35
53	99	67	140	56	261	42
54	147	50	211	33	369	18
55	87	71	138	56	272	40
56	58	80	110	65	189	58
57	202	32	234	26	379	16
58	118	60	198	37	318	29
59	9	97	8	97	34	92
60	20	93	85	73	234	48

TABLE 6:18(d)

Niagara Mathematics Test Series (Fundamental Operations), Grades V to VII (H. and M. Revision), 1976. Number of Correct Responses and Percentage of Incorrect Responses for each Question. Grade 5: n = 296; Grade 6: n = 315; Grade 7: n = 451.

Question No.	Grade 5		Grade 6		Grade 7	
	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect	# of Correct Responses	% Incorrect
61	7	98	52	83	153	66
62	64	78	169	46	359	20
63	77	74	135	57	247	45
64	118	60	173	45	350	22
65	91	69	160	49	318	29
66	131	56	185	41	354	22
67	32	89	92	71	164	64
68	39	87	89	72	229	49
69	32	89	52	83	110	76
70	29	90	51	84	149	67
71	11	96	76	76	255	43
72	22	93	99	69	317	30
73	19	94	13	96	141	69
74	45	85	108	66	270	40
75	61	79	105	67	207	54
76	29	90	52	83	113	75
77	143	52	199	37	380	16
78	116	61	193	39	354	22
79	97	67	134	57	238	47
80	23	92	36	89	97	78

TABLE 6:19(a)

Percentage of Students Missing Questions in Categories of Questions for the Niagara Mathematics Test Series, Fundamental Operations, Grades V-VII (H. and M. Revision), 1976.

Category of Questions	Percentage of Students Missing Questions		
	Grade 5 n = 296	Grade 6 n = 315	Grade 7 n = 451
1. Primary Operations			
(a) Addition:			
#1	2	2	1
#2	5	4	2
#6	7	6	6
#9	8	4	6
#12	26	20	14
Total:	9.6	7.2	5.8
(b) Subtraction:			
#3	12	9	5
#7	14	8	3
#10	23	13	13
#15	26	14	11
#18	30	22	12
Total:	21.0	13.2	8.8
(c) Multiplication:			
#4	16	12	7
#8	31	22	13
#11	39	30	17
#14	53	39	29
#17	60	46	33
Total:	39.8	29.8	19.8
(d) Division:			
#5	8	4	2
#13	16	12	15
#16	65	42	28
#19	73	56	54
#51	75	67	44
#53	67	56	42
#58	60	37	29
#73	94	96	69
#80	92	89	78
Total:	61.1	51.0	40.1

TABLE 6:19(b)

Percentage of Students Missing Questions in Categories of Questions for the Niagara Mathematics Test Series, Fundamental Operations, Grades V-VII (H. and M. Revision), 1976.

Category of Questions	Percentage of Students Missing Questions		
	Grade 5 n = 296	Grade 6 n = 315	Grade 7 n = 451
2. Problem-Solving:			
#41	81	76	66
#42	70	67	55
#43	69	64	53
#44	50	45	18
#54	50	33	18
#55	71	56	40
#56	80	65	58
Total:	67.3	58.0	44.0
3. Operational Definitions:			
#49	47	32	16
4. Weights and Measures:			
#21	60	42	27
#35	39	31	22
#40	63	43	24
#52	79	60	35
#66	56	41	22
#75	79	67	54
#76	90	83	75
#77	52	37	16
#78	61	39	22
#79	67	57	47
Total:	64.6	50.0	34.4
5. Fractions			
(a) Addition:			
#22	80	71	45
#23	84	74	46
#27	75	64	43
#33	81	60	37
Total:	80.0	67.3	42.8
(b) Subtraction:			
#24	59	56	35
#26	79	74	54
#57	32	26	16
#59	97	97	92
Total:	66.8	63.3	49.3

TABLE 6:19(c)

Percentage of Students Missing Questions in Categories of Questions for the Niagara Mathematics Test Series, Fundamental Operations, Grades V-VII (H. and M. Revision), 1976.

Category of Questions	Percentage of Students Missing Questions		
	Grade 5 n = 296	Grade 6 n = 315	Grade 7 n = 451
5. Fractions			
(c) Multiplication:			
#20	88	76	44
#25	92	74	51
#60	93	73	48
#61	98	83	66
Total:	92.8	76.5	52.3
(d) Division:			
#37	71	49	30
(e) Fractions (Other):			
#32	84	69	31
#63	74	57	45
#64	60	45	22
#65	69	49	29
Total:	71.8	55.0	31.8
6. Percentages:			
#70	90	84	67
7. Decimals			
(a) Addition:			
#28	75	58	35
#30	94	67	37
#34	87	49	16
Total:	85.3	58.0	29.3
(b) Subtraction:			
#29	60	45	26
#36	88	49	21
#39	86	50	23
#72	93	69	30
Total:	81.8	53.3	25.0

TABLE 6:19(d)

Percentage of Students Missing Questions in Categories of Questions for the Niagara Mathematics Test Series, Fundamental Operations, Grades V-VII (H. and M. Revision), 1976.

Category of Questions	Percentage of Students Missing Questions		
	Grade 5 n = 296	Grade 6 n = 315	Grade 7 n = 451
7. Decimals			
(c) Multiplication:			
#31	**89**	57	29
#38	**81**	49	29
#62	**78**	46	20
#71	**96**	76	43
Total:	**86.0**	57.0	30.25
(d) Decimals and Fractions:			
#68	**87**	72	49
(e) Decimals and Percentages:			
#69	89	83	76
(f) Decimals (Other):			
#67	89	71	64
#74	85	66	40
Total:	87.0	68.5	52.0
8. Factors:			
#45	33	22	10
#46	37	22	14
#47	58	52	29
#48	62	48	27
#50	86	72	53
Total:	55.2	43.2	26.6

7 Comparative Achievement in Language Arts

Table 7.1 shows the set of tests in the language arts, as listed in Table 3:3, which were employed in the 1975-76 phase of the study. As shown, these tests (except the Niagara Revision of the Vocabulary Test) were also administered in the 1950's phase of the study; in the 1930's phase either the results of the original tests or the Ontario norms were employed as a data base. It will be noted from the table that grade 8 students in 1975-76 wrote four tests, grade 6 students three tests, and grades 5 and 7 students two tests each.

Despite the number of tests employed, and the use of the general term "language arts" to describe these tests collectively, it would be improper to imply that this set of tests covered the total range of skills in a modern language arts program adequately. In fact, they evaluated only two main areas of language arts, silent reading with its subskills of word recognition, detail, inference, logic, and interpretation, etc., and student-generated vocabulary. Among others, the tests failed to deal with such areas of language arts as listening skills, discussion, creative writing, speaking, memory work, etc. The reader is cautioned therefore to avoid generalizing from the findings reported in this chapter to the broader set of language arts skills.

Some reference was made in Chapter 3 to the quality of the respective tests. In summary, it should be pointed out that the Thorndike-McCall Test (Grade 6 Paragraph Reading) was considered to be an excellent test in all three phases of the study. The two grade 8 reading tests were considered to be adequate, but changes in school objectives and vocabulary have lessened their applicability in the 1970's. The third test, the Vocabulary Test for Grades 5 to 8, was considered highly unsatisfactory for the 1970's and has therefore undergone, for the current study, more revision than any of the other tests, to the point that it was considered to be a new test and administered as such.

TABLE 7:1

Summary of Tests Used For Comparison of Student Achievement in Language Arts 1933-1938, 1952-1954, 1975-1976. *Refers to Ontario Urban Norms.

Name of Test	Grade(s) Tested	Date of Testing			Number of Pupils		
		1933-1938	1952-1954	1975-1976	1933-1938	1952-1954	1975-1976
Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Form A	8	June 1934	June 1952	June 1975	1000*	346	425
Dominion Group Achievement Test, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Form B	8	June 1934	June 1953	June 1975	1000*	340	410
Thorndike-McCall Reading Scale, 1920 Edition, Form 2	6	November 1938	November 1953	November-December 1975	370	441	280
Vocabulary Test Number 1, Ninety Common Names	5,6,7,8	November 1933	November 1953	January 1976	1648	1729	1672
Vocabulary Test Number 1, Ninety Common Names (Niagara Revision), 1976	5,6,7,8	----	----	April-May 1976	-----	-----	1573

Further reference to the applicability of the various tests will be made in appropriate sections of this chapter.

GRADE 8 ACHIEVEMENT IN READING

The data for the 1975 phase of the study for grade 8 reading were obtained by administering the 1934 editions of The Dominion Group Achievement Test in Reading, Part I, Forms A and B. In the discussion of these tests in Chapter 3 the investigator stated that a match between the tests and courses of study indicated that the tests were equally suitable in the 1930's, 1950's, and 1970's. However, after administering and marking the tests the investigator had increasing doubts about the suitability of a small number of items, particularly where usage of certain words or phrases had altered in the meantime. Despite these relatively minor considerations, the investigator continues to regard the tests as useful for basic comparison across the three phases of the study.

It will be recalled that the data from the 1930's were based on Ontario norms for 1934. Because there were no recorded figures for grade 8 students in St. Catharines on these tests in 1934 Partlow considered the urban norms for Ontario to be a suitable replacement. In comparing the norms from 1934 with the data in 1952-53, Partlow was confident in stating that there was an improvement in paragraph reading. His conclusion was phrased as follows:

> It appears that the less select group of grade 8's in 1952 and 1953 have achieved as high a standard or higher than the comparable group of grade 8 students would have achieved in 1934.

Table 7:2 and 7:3 compare the mean raw scores in 1975 with their counterparts in 1952-53 and the Ontario norms of 1934. The latter table also shows mean raw scores, adjusted in 1975 for socio-economic status and age-grade factors. Both adjustments were discussed at length in Chapter 4. Table 7:4 presents the data on the mean scores with accompanying standard deviations for

TABLE 7:2

Ontario Norms (1934) and Summary Data for Grade 8 Classes, St. Catharines Public Schools, 1952, 1975 from the Dominion Group Achievement Test, Part 1, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B.

Dominion Paragraph Reading Test	1934 Ontario Norms	1952		1975	
		Number of Pupils	Mean Raw Score	Number of Pupils	Mean Raw Score
Form A (Maximum Raw Score = 20)	14.0	346	14.9	425	14.24
Percentage of Maximum Raw Score	70.0%		74.5%		71.2%
Form B (Maximum Raw Score = 20)	14.0	340	14.0	410	14.36
Percentage of Maximum Raw Score	70.0%		70.0%		71.8%

TABLE 7:3

A Summary of Overall Means (With Adjustments for Socio-Economic Status and Age-Grade Level, 1975) for Grade 8 Students, St. Catharines Public Schools, 1934, 1953 and 1975, from The Dominion Group Achievement Tests, Part I, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B.

Test	1934 (Ontario Norms)	1952, 1953 Mean Raw Score	1975 Unadjusted Mean Raw Score	Adjustment For Socio-Economic Status, 1975	Adjustment For Age-Grade level, 1975	1975 Adjusted Mean Raw Score
Form A	14.0	14.9	14.24	+ 0.04	- 0.50	13.78
Form B	14.0	14.0	14.36	+ 0.78	- 0.51	14.63

TABLE 7:4

Class Means and Standard Deviations For Grade 8 Classes, St. Catharines Public Schools, 1975, from the Dominion Group Achievement Test, Part 1, Paragraph Reading, Grade VIII, 1934 Edition, Forms A and B.

School #	Form A		Form B	
	Class Means	Standard Deviations	Class Means	Standard Deviations
4	11.23	4.59	-----	-----
7	-----	-----	12.8	4.71
10	13.70	3.65	12.36	4.11
11	13.54 15.45	3.28 3.26	15.19 14.06	2.93 3.65
12	14.68 15.13 15.45	3.43 2.41 2.91	13.96 15.25 14.65 15.42 17.87	3.84 2.65 3.37 3.22 2.35

the various grade 8 classes writing each test in 1975. Unfortunately comparable data for class means and medians on these two tests are unavailable for the 1950's, and therefore a comparison cannot be made on the distribution of class means. However this table is useful not only to show the range of means across the study schools, but also as a base for similar comparisons in possible future phases of this longitudinal study.

Returning to Table 7:2, the reader notes a somewhat different pattern of comparative results for Form A and Form B of the test. On Form A of the test, the students in 1975, as shown by both the unadjusted and adjusted mean raw scores, performed slightly less well than their counterparts in the two earlier phases. On Form B, however, students in 1975 outperformed their counterparts in the two earlier phases.

The different comparative patterns for the two forms of the Dominion Group Achievement Tests of Paragraph Reading present a number of questions for which the data themselves do not provide adequate answers. For instance, one notes that the Ontario norms for both forms of the test are 14.0 in 1934. There are mean raw score differences, however, in 1952 and 1975. In these two specific years, the mean raw scores differ by approximately one raw-score point.

Similarly, one may wonder why the scores in Form A have risen from the 1934 norms in 1952 and then fallen below the 1934 norms in 1975 while the scores on Form B remained essentially the same in 1934 and 1952 and then increased in 1975 beyond mean raw scores for the two earlier years. Observation of Table 7:4 shows that when the performance of individual classes is inspected, the actual range, for 1975, of class means is from 11.23 to 15.45 on Form A and 12.80 to 17.87 on Form B. Clearly, these data suggest that while on the average grade 8 students in 1975 do slightly less well on Form A and slightly better on Form B than in the previous phases, there are individual classes in 1975 that significantly outperform average student performance in 1934 and 1954.

Is there a possibility, then, that either the 1952 or 1975 mean raw score is moderated by some factor unaccounted for in the research design? An examination of Form A (see appendices) reveals a number of items for which the vocabulary appears unsuited to the modern idiom. However, would this language match not have been reasonably unsuited to the students in 1954 as well? Obviously, the two Dominion Tests of Paragraph Reading cannot be considered, on the basis of these current observations, as parallel tests for either 1954 or 1975.

A further inspection of results on these two forms indicates that differences in 1975 attributable to the socio-economic factor were greater for Form B than for Form A. To satisfy the conditions required of parallel forms of a test, the effects of the socio-economic factor on student performance would have to be reasonably consistent. Similarly, the balance in language usage and its respective match with current vocabulary would also have to be equated. Nevertheless, since there are relatively few data to support the present investigator's conjecture about vocabulary and socio-economic changes and their differential importance in the two forms of the test, one is left without adequate explanation for different patterns in student performance on Forms A and B of the test.

In summary, the data obtained by combining the two sets of results indicate that the grade 8 students in the study schools today perform approximately as well as did students in the 1930's and 1950's. On the test (Form B) which showed the most consistent pattern over the period, and was judged preferable on other grounds, the students in 1975 performed slightly better than their earlier counterparts, both before and after adjustment for major variables. In attempting to account for the differences in these three periods, one might begin, as suggested above, with a more comprehensive analysis of the test content and its match with modern idiom and modern school objectives. Such an analysis might indicate that neither test measures adequately the broad array of content and skills that are now embodied in reading curricula. To support this suggestion, the reader is encouraged to compare the grade 8 results with those of grade 6 students (see Table 7:5) which show that grade 6 students in 1975 also

TABLE 7:5

Overall Distribution of Grade 6 Scores for St. Catharines Public Schools (November - 1938, 1953, 1971 and December - 1975) from the Thorndike-McCall Reading Scale, Form 2.

Crude Score	Grade Score	Number of Students				Per Cents			
		1938	1953	1971	1975	1938	1953	1971	1975
35									
34									
33	15.4								
32	15.2				1				0.36
31	14.9		1	1	2		0.23	0.33	0.714
30	14.6		2	1	5		0.45	0.33	1.79
29	13.8	1	3	5	13	0.27	0.68	1.67	4.64
28	12.6	2	9	11	13	0.54	2.04	3.67	4.64
27	10.5	7	10	17	19	1.89	2.27	5.67	6.79
26	8.8	11	20	22	20	2.97	4.54	7.33	7.14
25	8.0	26	34	40	27	7.03	7.71	13.33	9.64
24	7.4	36	49	24	26	9.73	11.11	8.00	9.29
23	6.8	40	57	37	26	10.81	12.93	12.33	9.29
22	6.5	50	62	31	24	13.52	14.06	10.33	8.57
21	6.1	37	55	41	25	10.00	12.47	13.67	8.93
20	5.8	50	49	18	21	13.52	11.11	6.00	7.5
19	5.4	37	33	17	16	10.00	7.48	5.67	5.71
18	5.1	18	28	11	13	4.86	6.35	3.67	4.64
17	4.9	24	9	10	12	6.49	2.04	3.33	4.29
16	4.5	12	10	5	7	3.24	2.27	1.67	2.5
15	4.3	10	5	4	4	2.70	1.13	1.33	1.43
14	4.2	5	3	2		1.35	0.68	0.67	
13	4.0	3	2	1	2	0.81	0.45	0.33	0.71
12	3.8	1			1	0.27			0.36
11	3.7			1	2			0.33	0.71
10	3.5			1				0.33	
9					1				0.36
Total		370	441	300	280				
Mean Crude Score		20.99	21.87	22.56	22.75				
Grade Equivalent		6.10	6.45	6.67	6.72				

outperform, on the average, their counterparts in 1934 and 1954 on a test which treats a somewhat broader set of reading skills.

Other possible reasons for the superior scores of grade 8 students on Form A in 1954 as compared to the 1970's were given in the previous chapter. The presence in modern grade 8 classrooms of students who in the earlier phases would have been retained in lower grades or would have left school is one obvious explanation. Others include the different emphasis on the particular reading skills under study, limitations now placed on language arts programs by rotary time-tables and by competition with other subjects, and the reduced use of formal examinations. Together, these differences may account for the slight apparent drop in student performance on one form of the test in paragraph reading at the 8th grade level. In assessing these remarks about the drop in scores on Form A, the reader is reminded that the current grade 8 performance on Form B exceeded that of the earlier groups.

GRADE 6 ACHIEVEMENT IN READING

Of all the tests used in either subject area in the three phases of the longitudinal evaluation, none offers as suitable and consistent a basis for comparison of achievement across the intervening years as the Thorndike-McCall Reading Scale. Although it was constructed in 1920 and has a number of weaknesses, as discussed in Chapter 3, its form, construction, organization, and content make it still a highly suitable test of grade 6 reading, and almost ideal for utilization in a longitudinal study of this kind. As stated earlier it has by far the highest validity of any of the original tests. The organization of the test, ranging from very simple and easy to highly complex and difficult items, with the easiest at the beginning and the most difficult at the end, makes it highly suitable for comparing growth in student achievement. The aspects of the test that made it slightly less than ideal were relatively few in number, and involved mainly changes in vocabulary or current meanings of words, some of which were discussed in Chapter 3; in addition, the range of acceptable answers in the teachers' manual in 1920 indicated a measure of conformity or arbitrariness in the selection of single answers from among what is a set of equally acceptable answers today. Despite these minor problems this test remains highly suitable for all phases of the study.

Obviously a similar observation was made by Dr. Partlow, who reported that this test was "the most useful of the reading tests for the study as it is a carefully standardized test which was given to all St. Catharines grade 6 students in November, 1938 and November, 1953."

Not only was this test itself superior to any other, but the available data on this test were also superior. The investigators were able to administer exactly the same test in all three phases. Furthermore, complete data could be accurately retrieved from the tables from the 1930's and 1950's. Mean crude scores, standard deviations, and grade equivalents were computed and tabled for inspection.

One further asset accompanies these data. Donald Edward Hastings, a principal in the school system, administered this test to the same schools in 1971, and thereby has provided the study with data for an intermediate year between the second and third phases of the study.

Table 7:5 shows the number and percentage of students scoring at each level in the 1938 and 1953 phases of the study and in the two phases in the 1970's, namely the Hastings study in 1971 and the 1975 phase of the current study. This table shows that the mean crude scores in both studies in the 1970's were higher than their counterparts in the 1930's and 1950's. The grade equivalents, based on mean crude scores in each case, were 6.10 in 1938; 6.45 in 1953; 6.67 in 1971; and 6.72 in 1975. The data indicate that there has been a reasonably consistent gain between 1938 and the 1970's, with the two studies in the 1970's reporting fairly similar mean crude scores and grade equivalents.

The table also indicates that there is a much greater variation in student achievement levels in the 1970's than in either of the first two phases of the study. That is, the range of student performance in 1975 is greater than in the earlier two phases. This finding supports the investigator's observation that classes today are more heterogeneous than in the past. Despite there being a relatively large number of low achievers in reading, the table shows that the average achievement level in both 1971 and 1975 is higher than in 1938 and 1953. In both 1971 and 1975 the highest achievers on the Thorndike-McCall Reading Scale outperformed the highest achievers in either of the previous studies.

Tables 7:6(a), 7:6(b), and 7:6(c) give a school-by-school breakdown of the grade 6 students' performance on the Thorndike-McCall Test for the three major phases. Observation of these tables shows that in every instance in 1975 grade 6 pupils outperformed, on class averages, their counterparts in the two earlier periods. Hastings found that the median grade scores in 1971 were higher than their earlier counterparts in five schools and were lower in two schools. It is noteworthy that the two schools in which Hastings reported a drop in grade 6 medians from 1953 to 1971 were reported in the current study as having faced a relative socio-economic drop in the intervening years.

The overall summary of the raw score means for grade 6 reading across the four applications of this test are shown in Table 7:7 and in Figure 7:1. Here the raw figures for 1971 and 1975 are given, along with adjustments for the 1975 age-grade and socio-economic factors as developed and described in Chapter 4. In addition the results from the 1971 study are adjusted by the same amounts as the 1975 study, on the assumption that these two factors would have applied in the same directions and to approximately the same extent in 1971 as in 1975. It is even reasonable that the 1971 researcher did not comment on these two factors or speculate on their extent, because of the fact that his was a relatively small-scale project in comparison to the present study. In order to make comparisons on the same central measures, the raw data from the Hastings study, which like the earlier studies employed medians, were recalculated to produce means; thus

TABLE 7:6(a)

Distribution of Grade 6 Scores from The Thorndike-McCall Reading Scale, Form 2, by Schools and Classes, St. Catharines Public Schools (November - 1938, 1953 and December 1975), for Schools #1, #2, and #3.

Crude Score	Grade Score	SCHOOL #1					SCHOOL #2				SCHOOL #3*
		1938		1953		1975	1938	1953		1975	1938
		A	B	A	B			A	B		
35											
34											
33	15.4										
32	15.2										
31	14.9										
30	14.6			1		1			1	2	
29	13.8					2				1	
28	12.6					3			1	1	
27	10.5		1	3		1		1		2	1
26	8.8	3	1	1	1		2			2	1
25	8.0	1	2	3	4	7	3		3	7	3
24	7.4	2	5	4	3	5	2		2	3	5
23	6.8	1	3	6	2	5	1	2	8	6	6
22	6.5	5	7	3	9	5	5	2	6	2	2
21	6.1	4	8	2	5	4	3	2	6	1	7
20	5.8	6	3	5	5	2	1	3	5	4	3
19	5.4	5	2	3	2	1	7	4	3	5	3
18	5.1	4	1	2	1	1	2	1	1	4	2
17	4.9	1	2			2	2	2	1	2	2
16	4.5	1		3	1	2	3		1		1
15	4.3				3		1	1		3	1
14	4.2	1			1				1		2
13	4.0					1			1		
12	3.8										
11	3.7										
10	3.5										
9											
No. of Pupils		34	35	36	37	42	32	18	40	45	39
Mean Crude Score		20.65	21.83	22.08	21.03	22.83	20.53	20.11	21.30	22.20	21.10
Standard Deviation		2.84	2.33	1.40	1.62	3.81	3.15	2.74	3.22	3.98	3.58
Grade Equivalent		6.0	6.5	6.5	6.1	6.7	6.0	5.8	6.2	6.5	6.1

* School #3 was closed after 1938.

TABLE 7:6(b)

Distribution of Grade 6 Scores from The Thorndike-McCall Reading Scale, Form 2, by Schools and Classes, St. Catharines Public Schools (November - 1938, 1953 and December 1975), for Schools #4, #5, and #6.

Crude Score	Grade Score	SCHOOL #4				SCHOOL #5			SCHOOL #6		
		1938	1953		1975	1938	1953	1975	1938	1953	1975
			A	B							
35											
34											
33	15.4										
32	15.2										1
31	14.9							1		1	
30	14.6							1			
29	13.8		1		3		1	1	1		2
28	12.6						4	2		1	1
27	10.5		1		3	1	1	4	2		4
26	8.8		3	3	4		2	5		3	5
25	8.0		2	4	3	3	2		4	2	1
24	7.4	1	2	3	5	4	10	3	3	7	4
23	6.8	5	2	3	3	2	3	4	7	5	5
22	6.5	10	2	8	7	2	4	2	6	5	4
21	6.1	1	4	4	4	1	4	3	2	4	6
20	5.8	8	5	5	3	3	3	4	6	4	2
19	5.4	2	2	2		3	2	1	2	7	1
18	5.1	1	4	4	1		2	1	2		3
17	4.9	4	3	2	3		1	3	1		
16	4.5		1	2					3		1
15	4.3	1		1							
14	4.2	2		1							
13	4.0	2	1								1
12	3.8										1
11	3.7				1						
10	3.5										
9					1						
No. of Pupils		37	33	42	41	19	39	35	39	39	42
Mean Crude Score		19.81	21.09	21.05	22.51	22.42	23.13	23.51	21.77	22.49	22.86
Standard Deviation		3.06	3.62	3.13	4.25	2.43	3.02	3.84	3.08	2.76	4.13
Grade Equivalent		5.8	6.1	6.1	6.6	6.6	6.8	7.1	6.4	6.6	6.8

TABLE 7:6(c)

Distribution of Grade 6 Scores from The Thorndike-McCall Reading Scale, Form 2, by Schools and Classes, St. Catharines Public Schools (November - 1938, 1953 and December 1975), for Schools #7, #8, and #9.

Crude Score	Grade Score	SCHOOL #7					SCHOOL #8					SCHOOL #9		
		1938		1953		1975	1938		1953		1975	1953		1975
		A	B	A	B		A	B	A	B		A	B	
35														
34														
33	15.4													
32	15.2													
31	14.9					1								
30	14.6										1			
29	13.8										2		1	2
28	12.6	2				2			1	2	1			3
27	10.5		2	1					1	1	4	1		1
26	8.8	1	1	1	1	1		2	2	2	2	1		1
25	8.0	2	2	3	3			6	3	2	4	1	2	5
24	7.4	4	5	4	2	1	4	1	5	5	3		2	2
23	6.8	4	3	6	1	1	2	6	5	6		3	5	2
22	6.5	4	3	4	7		4	2	3	3	3	2	4	1
21	6.1	1	5	4	8	2	2	3	4	3		2	3	5
20	5.8	7	4	1	3	1	5	4	3	4	4		3	1
19	5.4	3	3	1	3	2	5	2	2	1	1		1	5
18	5.1		1	4	1	2	3	2	1	1			6	1
17	4.9	2	3			1	4	3			1			
16	4.5	1	1				2				1	1	1	3
15	4.3	2	2				1	2			1			
14	4.2													
13	4.0						1							
12	3.8						1							
11	3.7										1			
10	3.5													
9														
No. of Pupils		33	35	29	29	14	34	33	30	30	29	11	28	32
Mean Crude Score		21.39	21.2	22.24	21.69	22.36	19.44	21.39	22.77	22.97	23.17	22.64	21.29	22.36
Standard Deviation		3.32	3.24	2.45	2.0	4.41	3.07	3.19	2.47	2.55	4.60	2.94	2.86	3.89
Grade Equivalent		6.2	6.1	6.5	6.4	6.6	5.6	6.2	6.7	6.8	6.9	6.6	6.2	6.7

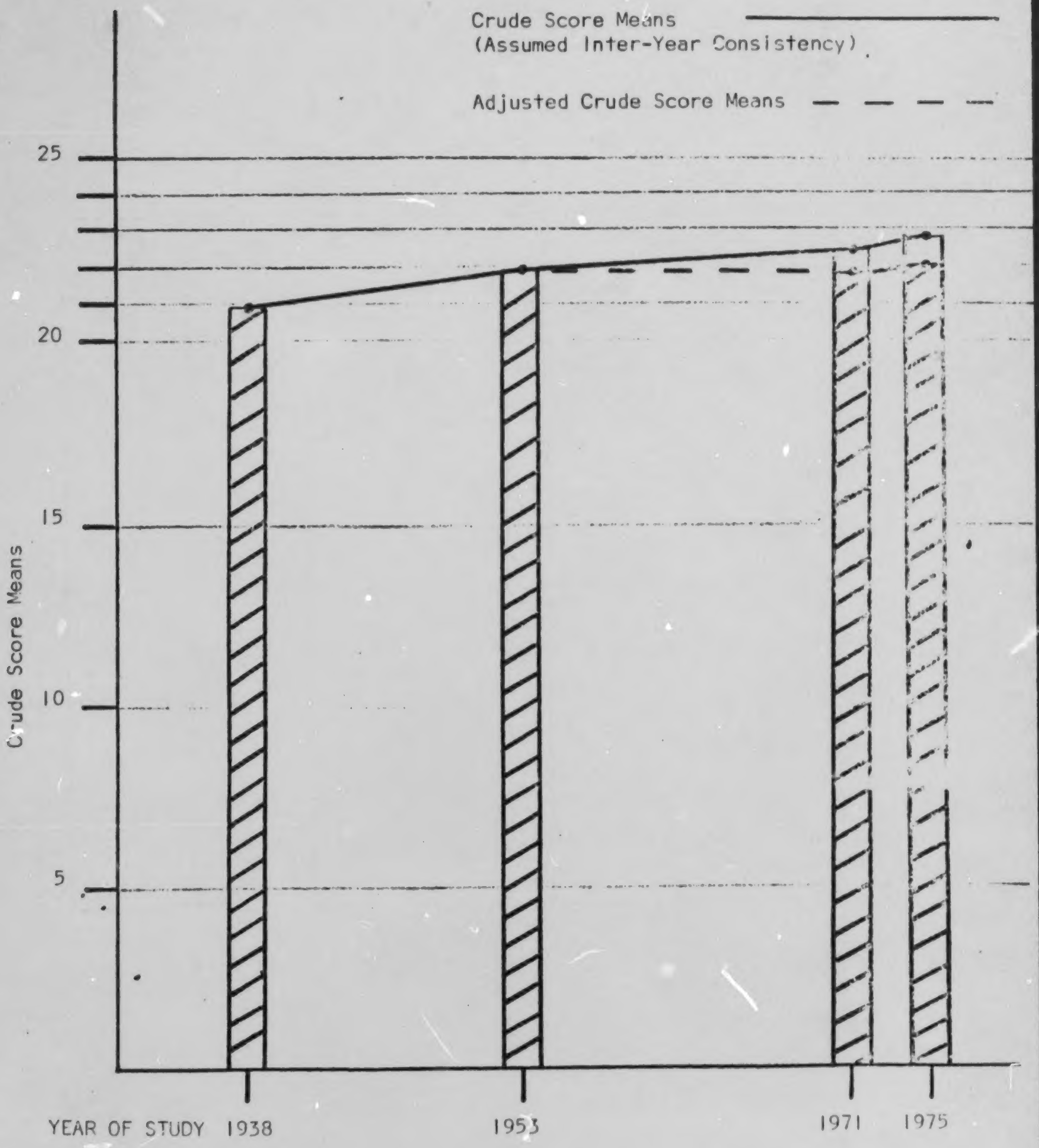
TABLE 7:7

Summary Data for Grade Six Classes, St. Catharines Public Schools, November - 1938, 1953, 1971 and December - 1975 (With Adjustments For the Age-Grade Differences in 1971, 1975), from the Thorndike - McCall Reading Scale, Form 2.

	1938	1953	1971	1975
Number Tested	370	441	300	280
Mean Crude Score	20.99	21.87	22.53	22.75
Age-Grade Adjustment	(none)	(none)	-0.75	-0.75
Adjusted Mean Crude Score	20.99	21.87	21.78	22.00
Standard Deviation	3.11	3.01	3.42	4.05
Mean Grade Score (From Unadjusted Means)	6.10	6.45	6.67	6.72

FIGURE 7:1

A Comparison of Grade Six Crude Score Means for St. Catharines Public Schools, 1938, 1953 1971, and 1975 (With Adjustments for Socio-Economic Status and Age-Grade for 1971 and 1975) Based on the Thorndike-McCall Reading Scale, 1920 Edition, Form 2.



the figures for both studies in the 1970's represent means.

Although both Hastings and the present investigator noted differences in achievement in different schools, there was no consistent pattern of relationships between socio-economic factors and differences in achievement, and therefore the socio-economic adjustment in this test is negligible. Such, however, is not the case with age-grade. Since the average grade 6 student in 1975 was approximately 0.4 years older than his counterparts a generation or two ago, an adjustment of -0.75 raw score points is made to the 1975 raw score means. As stated above, the same adjustment is applied to the 1971 figures.

In summary, an analysis of Tables 7.5 to 7.7 established that there is a slight but definite increase in reading achievement between 1938 and the present, and that the increase has apparently remained fairly steady from the 1930's into the mid-1970's. The two studies four years apart in the 1970's produced remarkably similar scores. Even after adjustments have been made for socio-economic factors and age-grade differences in the 1970's, the results still remain higher for students in the 1970's. Therefore the investigator concludes that students in the 1970's in the study schools in St. Catharines outperformed their counterparts of the same age in the 1930's and the 1950's by a narrow margin. The quality of the test involved, coupled with the breadth and quality of the data available in four different test periods with tests conducted and results analyzed by four independent researchers, adds credibility to these findings. In view of the validity and quality of all the factors involved, and because reading has been considered to be one of the more important objectives of the school throughout the period under study, the results of this component of the project can be considered to be the most precise indication of the overall comparative performance of students.

COMPARATIVE ACHIEVEMENT IN VOCABULARY SKILLS, GRADES 5 to 8

It will be recalled, from the discussion in Chapter 3, that the original Vocabulary Test of Common Names was administered in

1933, 1953 and 1975, and that the Niagara Revision of this test was administered in 1976. Also in Chapter 3, some of the weaknesses of the original test are discussed, along with the procedures for the preparation of the Niagara Revision.

While the Thorndike-McCall Reading Test described in the previous section was considered to be the best of all the tests employed, the original Vocabulary Test was considered for a number of reasons to be the least useful. First of all it measured only a very limited component of the language arts program - the student's ability to generate names of objects from their descriptions. This test was considered to have shortcomings when it was used in 1933, and in 1953 Partlow recognized that it had serious weaknesses in addition to those identified in 1933; in fact he regarded changes in student vocabulary to have occurred in the intervening period to the point that he applied a procedure to counteract the effects of these changes. By 1976 the test was so unsuitable that entire classes in the preliminary test run attained zero scores on a large number of items in the test.

Not only were many of the test items themselves entirely unsuitable in 1976, but also the marking scheme for many items was no longer satisfactory. Changes in both the meaning and usage of words, and striking changes in the range of correct answers given by students in response to descriptions, had altered so much in the forty years since the original administration of the test that the marking scheme would have had to be almost completely revised even if the test itself were retained. For these and other reasons it was obvious that if a comparable test of vocabulary skill was to be given, it would have to be a new test, employing as many of the existing items as were still suitable, applying the same format, and replacing obsolete items with items equal in difficulty to those removed from the test. Despite the large amount of work that went into designing the revised test, and the diligent attention that a large number of capable referees gave to the task of generating, revising, and validating test items, the investigator is less than completely satisfied with the resulting instrument, particularly in respect to its comparability to the earlier test. Without better information on the frequency of use by children of words in the original test at the time it was constructed, the referees found it extremely difficult to determine the comparative difficulty of original items and items in the Niagara Revision. As a result, despite efforts to match difficulty, there remains a possibility that the Niagara Revision is somewhat easier than the original test.

In the paragraphs that follow, the results, first for the original test in 1975, and then for the Niagara Revision, are presented and compared. Table 7:8 shows, for the original test (Vocabulary Test Number 1), the medians and percentages for 1933 and 1953, and the means and percentages for 1975. It will be noted that in all grades the mean scores in 1975 dropped below both the 1933 and 1953 medians; the drop was from about one to five percentage points across the four grades. The change from medians to means as measures of central tendency accounts for a portion of this drop, but the greater portion represents a consistently lower level of performance on this test. In view of the weaknesses of the original test, as described earlier, as an instrument to measure vocabulary skills in the 1970's, the scores are surprisingly high, suggesting the possibility that on the items that have remained valid the students today outperform their earlier counterparts. This possibility appears to be sustained by the evidence that follows, concerning achievement on the Niagara Revision of the test.

Tables 7:9(a) to 7:9(d) show the ranked medians for 1933 to 1953 on the original test, along with the ranked means for 1976 on the Niagara Revision for each of the classes in the study. It should be noted that the means presented for 1976 in these tables are not adjusted at this point for either socio-economic status or age-grade changes; these adjustments appear in Table 7:10 where the data for all the grades are presented for comparison. Nevertheless, Tables 7:9(a) to 7:9(d) present considerable information about the comparative performance of the various grades in the three test

TABLE 7:8

A Comparison of Grade 5, 6, 7, 8 Student Performance, St. Catharines Public Schools, Based on Vocabulary Test Number I, Ninety Common Names, 1933, 1953 and 1975.

	1933		1953		1975	
Grade	Median	Percentage Score	Median	Percentage Score	Mean	Percentage Score
5	27	30.0%	28	31.1%	27.1	30.1%
6	37	41.1%	35	38.9%	30.6	34.0%
7	44	48.9%	42	46.7%	37.5	41.7%
8	52	57.8%	45	50.0%	41.6	46.2%

TABLE 7:9(a)

A Comparison of Grade 5 Class Medians, Vocabulary Test Number One, 1933 and 1953, and Class Means, Vocabulary Test of Common Names (Niagara Revision), 1976, for St. Catharines Public Schools.

Ranked Class Medians		Ranked Class Means
1933	1953	1976
36	37	63.50
33	32	62.71
31	30.5	55.92
27	29.5	55.10
26	29.5	54.7
26	29	52.8
26	28	52.1
25	28	50.90
25	27	50.89
25	27	50.41
24	26	49.60
22	25.5	47.91
17	25	47.20
	24.5	31.82
	21	
City Median 27	City Median 28	City Mean 51.79 (S.D. 14.77)

TABLE 7:9(b)

A Comparison of Grade 6 Class Medians, Vocabulary Test Number One, 1933 and 1953, and Class Means, Vocabulary Test of Common Names (Niagara Revision), 1976, for St. Catharines Public Schools.

Ranked Class Medians		Ranked Class Means
1933	1953	1976
	47	65.70
44	39	64.00
43	39	62.19
43	36	61.08
38	36	60.60
38	36	59.27
36	35	57.80
36	33.5	57.51
35	33	55.88
35	33	51.67
35	32	50.75
34	32	50.75
32	31	49.70
	29.5	48.80
City Median 37	City Median 35	City Mean 57.88 (S.D.13.48)

TABLE 7:9(c)

A Comparison of Grade 7 Class Medians, Vocabulary Test Number One, 1933 and 1953, and Class Means, Vocabulary Test of Common Names (Niagara Revision), 1976, for St. Catharines Public Schools

Ranked Class Medians		Ranked Class Means
1933	1953	1976
49	55	76.80
47	48	75.30
47	47	74.20
46	46	71.60
46	46	71.40
45	44	67.15
45	43	65.70
43	41	65.25
40	40.5	64.72
39	39	64.72
36	39	64.70
	37.8	64.70
	35	63.80
		63.20
		62.10
		57.74
		51.60
City Median 44	City Median 42	City Mean 66.26 (S.D. 12.52)

TABLE 7:9(d)

A Comparison of Grade 8 Class Medians, Vocabulary Test Number One, 1933 and 1953, and Class Means, Vocabulary Test of Common Names (Niagara Revision), 1976, for St. Catharines Public Schools.

Ranked Class Medians		Ranked Class Means
1933	1953	1976
		79.00
		77.10
55	53	76.22
55	48	75.40
55	48	75.10
54	48	75.00
51	46	72.15
51	45	72.00
47	44.5	71.33
46	44	68.83
	43	68.33
	41.5	68.32
	39	68.00
		67.81
		67.70
		67.20
		66.95
		63.89
		46.84
City Median 52	City Median 45	City Mean 70.39 (S.D. 11.56)

TABLE 7:10(a)

Class Means and Standard Deviations, Grade 5 Students, St. Catharines Public Schools, on the Vocabulary Test of Common Names (Niagara Revision), 1976.

School	# of Students	Mean	Standard Deviation
1	29	50.89	14.42
1	25	55.92	11.69
2	27	49.60	14.07
2	28	52.10	16.92
4	12	50.41	12.82
4	27	54.70	11.46
5	21	63.50	12.80
5	16	55.10	15.91
6	25	50.90	13.34
6	7	62.71	7.48
7	11	31.82	9.50
8	25	47.20	14.20
9	16	52.80	20.26
9	32	47.91	12.15

TABLE 7:10(b)

Class Means and Standard Deviations, Grade 6 Students, St. Catharines Public Schools, on the Vocabulary Test of Common Names (Niagara Revision), 1976.

School	# of Students	Mean	Standard Deviation
1	29	59.27	11.47
1	26	55.88	12.88
2	21	62.19	10.17
2	23	60.60	13.05
4	25	48.80	18.25
4	25	57.80	13.26
5	25	64.0	9.15
5	10	49.70	10.21
6	23	65.70	12.41
6	17	51.47	14.44
7	12	50.75	12.98
8	33	57.51	15.00
9	25	61.08	10.64
9	6	51.67	9.77

TABLE 7:10(c)

Class Means and Standard Deviations, Grade 7 Students, St. Catharines Public Schools, on the Vocabulary Test of Common Names (Niagara Revision), 1976.

School	# of Students	Mean	Standard Deviation
4	22	64.72	14.92
7	14	75.30	12.83
9	22	64.70	9.49
10	27	63.80	12.04
10	29	65.70	10.81
10	29	64.70	11.88
10	25	63.20	17.43
11	29	62.10	12.26
11	29	64.72	10.17
11	18	51.60	17.08
11	31	71.60	8.74
11	27	57.74	10.07
12	27	67.15	8.13
12	27	65.25	11.22
12	28	74.20	6.84
12	28	76.80	5.79
12	25	71.40	10.45

TABLE 7:10(d)

Class Means and Standard Deviations, Grade 8 Students, St. Catharines Public Schools, on the Vocabulary Test of Common Names (Niagara Revision), 1976.

School	# of Students	Mean	Standard Deviation
4	59	67.20	16.11
7	16	75.00	11.75
9	24	68.00	6.82
10	30	71.33	11.66
10	27	67.70	11.42
10	30	68.83	8.65
10	28	63.89	12.56
11	19	46.84	7.75
11	32	67.81	9.59
11	34	72.15	10.53
11	21	66.95	11.60
11	31	76.22	5.19
11	32	79.00	5.86
12	28	75.40	6.56
12	28	68.32	7.26
12	32	72.00	6.42
12	34	77.10	7.38
12	30	75.10	6.77

periods, and also about the relative heterogeneity of the classes in 1976. The reader will note that in almost every case the means in 1976 are higher than the medians in any of the previous years. The range of scores and standard deviations from class to class and school to school, shown in Tables 7:10(a) to 7:10(d) and summarized in Table 7:11, provides further evidence of the greater heterogeneity of classes today, an observation referred to earlier in this report.

The overall comparison of the data is provided in Table 7:12 and Table 7:13. In the latter table, the 1976 means are adjusted for socio-economic and age-grade differences as outlined in Chapter 4. An inspection of this table establishes the apparent superiority of the 1976 students in the study schools relative to their counterparts in 1933 and 1953. However, the size of the increases in these measures of central tendency requires additional comment; in fact, one wonders on first inspection whether these increases are even plausible. In the attempt to assess these differences, a number of factors merit consideration. Elsewhere in this report a comparison of the use of medians in the earlier phases and means in the present phase has been explained. It is unlikely that the use of medians for the 1976 group would decrease by any significant amount the large difference in the scores. The investigator proposes that the vocabulary test used in the original studies may, in fact, have been less suited to the populations in the first two phases of the study than either of the original authors perceived. The study of the median scores for the four grades in question in both 1933 and 1953 lends substance to this argument. In these two phases grade 5 students had only 30 percent and 31 percent of the answers correct; grade 6 had 41 percent and 39 percent correct; grade 7 had 49 percent and 47 percent correct; and grade 8 had 58 percent and 50 percent correct. A test designed for grades 5 to 8, on which the top mean score for grade 8 is only 58 percent does not provide a great enough range to assess growth from grade to grade adequately, and is too difficult for any of the grades. No standardized data could be found for the instrument during the

TABLE 7:11

Overall Means and Standard Deviations, Grades 5, 6, 7, and 8, for St. Catharines Public Schools, on the Vocabulary Test of Common Names (Niagara Revision), 1976.

Grade	Number of Classes	Overall Mean	Standard Deviation
5	14	51.79	14.77
6	14	57.88	13.48
7	17	66.26	12.52
8	18	70.39	11.56

TABLE 7:12

Comparison of Medians, Means, and Percentage Scores for Grades 5, 6, 7, and 8, from the Vocabulary Test Number 1, Ninety Common Names, for 1933, 1953, and 1975, and the Vocabulary Test of Common Names (Niagara Revision), 1976.

	1933		1953		1975			1976		
Grade	Median Score	% Score	Median Score	% Score	No. of Pupils	Mean Score	% Score	No. of Pupils	Mean Score	% Score
5	27	30.0	28	31.1	296	27.10	30.11	302	51.79	57.54
6	37	41.1	35	38.9	287	30.69	34.10	299	57.88	64.31
7	44	48.9	42	46.7	558	37.50	41.67	437	66.26	73.62
8	52	57.8	45	50.0	531	41.61	46.23	535	70.39	78.21
Total Number of Students	1648		1729		1672			1573		

TABLE 7:13

A Summary of Medians and Means (With Adjustments for Socio-Economic Status and Age-Grade Level, 1976) for Grade 5, 6, 7, 8 Students, St. Catharines Public Schools, from the Vocabulary Test Number 1, Ninety Common Names, 1933, 1953 and the Vocabulary Test Number 1, Ninety Common Names (Niagara Revision), 1976.

Grade	1933 Medians	1953 Medians	1976 Unadjusted Raw Score Means	1976 Adjustment for Socio-Economic Status	1976 Adjustment for Age-Grade	1976 Adjusted Raw Score Means
5	27.00	28.00	51.79	+3.68	-2.10	53.37
6	37.00	35.00	57.88	+1.73	-2.80	56.81
7	44.00	42.00	66.26	+0.38	-2.10	64.54
8	52.00	45.00	70.39	+1.50	-3.50	68.39

period when it was in use in England in the early 1930's, and in the face of this lack of normative information the investigator assesses the instrument as having been less valid even during the 1930's and 1950's than was assumed by the two original investigators. The drop in scores from 1933 to 1953 suggests that many of the test items were rapidly becoming obsolete during the intervening two decades. On inspection of the test, the investigator challenged the suitability of some of the questions even for the time when the test was originally administered, and on administering the test to adults who had lived during that decade he found that numerous terms baffled them, indicating that these words probably were not in common use in Ontario even at that time.

The range of mean percentages on the 1976 test as shown in Table 7:12, namely 57 percent for grade 5; 64 percent for grade 6; 74 percent for grade 7; and 78 percent for grade 8, represents roughly the range of scores for succeeding grades that would be expected on a test designed to measure growth in children's ability to generate vocabulary from definitions. However, these percentages may suggest that the modern tests produced higher scores for the converse of the previous argument, i.e., the 1976 test may have been easier. Since we have no method for determining what the achievement in 1933 and 1953 would have been on the 1976 test, we must allow for the possibility that the revised test was easier for students in 1976 then the original test was for students in the 1930's and 1950's. It is doubtful, however, that this possibility would account for the large extent of the difference in scores. The investigator concludes that students in these four grade levels are more proficient in this skill than their counterparts of one and two generations ago. Readers are cautioned, however, to apply to this general conclusion the restraints required by the use of an unstandardized test developed elsewhere for which only sketchy data were recorded and for which there was at best an uneven match with the curriculum of the time.

The test results showed another dimension of this skill not revealed by the earlier two studies, namely the tendency of students in 1976 to provide a much wider range of correct answers for many of the descriptions than apparently was the case in the 1930's and 1950's. For many of the test items, students provided three or four and in some cases up to six responses all of which were equally correct. The rather narrow set of acceptable answers for the original test items no doubt accounted for some of the depression of the scores in the 1930's and 1950's. In analyzing the acceptable answers to certain items even at that time, it appears that other equally correct answers, at least in the Canadian context, could have been and undoubtedly were suggested by students. It is not clear as to whether the decision at that time to accept only a very limited set of correct answers had to do with arbitrariness of those setting the marking scheme or reflected a much higher degree of conformity by teachers with respect to vocabulary than would be expected today. The investigator suggests that the deliberate encouragement by teachers to have students offer more creative answers to questions and to use a wider range of vocabulary has had a positive effect on the student's use of alternative words. Since this is a test of student-generated vocabulary, one would suppose that the tendency to produce a broader set of correct answers would be regarded as a positive attribute, and it was in this light that the marking scheme for the 1976 test recognized a broader set of correct answers.

While failing to produce unqualified conclusions concerning the comparative performance of students in the 1930's, 1950's, and 1970's with respect to their ability to generate vocabulary from descriptions, the revised test and the data stemming from it are highly useful for future studies, and further investigators are encouraged to use both the test and the data from 1976 as a valid indication of student performance at this time.

Given the general but somewhat qualified conclusion that students perform better in the skill of generating vocabulary than they did a generation or two ago, one must speculate further concerning some of the factors which may have encouraged this change. A wider range of books available to students; the domination of students' out-of-school time by television with its repetitious use of alternative vocabulary; the breadth and quality of discussion in classrooms; and the greater attention of teachers to creative expression appear to be factors which have enhanced the contemporary students' competency in this skill.

SUMMARY OF LANGUAGE ARTS ACHIEVEMENT DATA

In attempting to summarize the findings in the language arts area, the investigator finds there is strong evidence for concluding that students do as well or better than students in the 1930's and 1950's in the study schools in those areas of language arts under investigation. No data or inferences are available concerning other components of the language arts program. Also, the reader is cautioned against applying without careful consideration these generalizations, taken from studies in a few core schools in one city, to the population at large.

In the most respected test in the series, which also generated the most useful total set of data, the evidence is clear that students in the 1970's in grade 6 reading outperformed students in the 1930's and 1950's. This conclusion is supported by independent studies in the 1970's, both of which produced remarkably similar mean scores and grade equivalents.

Data obtained from the two forms of the 1934 Dominion Paragraph Reading Tests indicate that grade 8 students today perform approximately as well in paragraph reading as their earlier counterparts. This conclusion speaks well for this component, when one takes into account changes in the suitability of the tests, altered conditions in the populations of grade 8 classes, and the reduced emphasis given to the skill under study.

With respect to the skill of vocabulary use as reflected by a student's ability to generate a correct word from a given definition, the evidence indicates that students at all the grade levels under study performed better than their counterparts one and two generations ago and also showed greater range and variety in the alternative words provided. Some qualification must be made in this generalization in view of important differences which may have occurred with respect to the comparative suitability of the tests used across the various periods under study.

8 Conclusions

The preceding two chapters contain the comparative results across the three phases of this longitudinal study of student achievement in basic skills in mathematics and language arts. Each of the two basic skill areas was looked at separately on a grade-by-grade and test-by-test basis. Data were analyzed in light of the investigator's original intention to compare student performance in 1975-76 and in two previous periods, using tests of fair comparison with interpretation of the results underpinned by appropriate consideration of changes in four main variables - age-grade levels, socio-economic factors, vocabulary, and school objectives. In the final two chapters of the report the investigator intends to set forth the major conclusions of the study and to describe other outcomes of the project. In presenting the conclusions, the investigator will approach the findings in the same sequence in which they appear in the two previous chapters, and then will discuss conclusions concerning the effects of the four major variables under study and an array of other changes in schools which have certain effects on the data collected, but for which no specific measures of effect were attempted in this project.

In studying these conclusions and inferences, the reader is again reminded that the actual data pertain only to the four grades in question in a group of schools in that part of St. Catharines which formed the city proper in the 1930's and 1950's. Because of similarities in the history of schools in the urban areas in Ontario during the past forty years, the reader may feel justified in generalizing these conclusions beyond the study schools. Before making such generalizations, however, the reader is cautioned to consider the similarities and differences that may exist between the study schools and the populations to which the generalizations are extended.

ACHIEVEMENT IN MATHEMATICS

In Chapter 6 the investigator reported the data which led to the conclusion that between the 1950's and 1975 there has been a sizeable decrease in student performance in grade 8 in both arithmetic computation and arithmetic reasoning. The levels of student achievement in these two interrelated mathematics skills in grade 8 were below the levels in the 1930's and the 1950's, and even when adjusted for changes in socio-economic factors and age-grade levels, the performance in 1975 was considerably below that of the previous two periods. Particular attention should be given to the fact that of all the skills tested in both broad subject areas, and in all four grades tested, these two mathematics skills at the grade 8 level were the only two that showed any consistent loss in the 1970's phase of the study. Some of the apparent reasons for the drop in grade 8 achievement in these two areas of mathematics were presented in Chapter 6, and further consideration of these and a number of other factors are presented later in this chapter.

It must be emphasized that the decrease in achievement in mathematics computation and reasoning at the grade 8 level cannot be attributed only to changes in the four major variables under study. When a revised test considered to reflect modern content and objectives of grade 8 programs in these two skills areas was administered, and when the results were adjusted for socio-economic and grade-age factors, the results again showed that the students were consistently below those in the 1930's and 1950's in these skills.

A marked difference was found in the comparative achievement of grades 5, 6, and 7, as contrasted to that in grade 8. While the grade 5 to 7 classes performed slightly less well than students in the earlier two phases in most of the subskills in the original tests, when a test of fair comparison reflecting modern mathematics content and objectives was used the students in all three grades obtained higher scores than their counterparts on the original tests. The rationale supporting the assumption of fair comparison is presented in considerable detail elsewhere in this report. The results on this test, however, warrant restatement here. In general, students in grades 5 to 7 today outperform, on average, their counterparts in 1936 and 1954 in fundamental operations in mathematics. Even after scaling downward the weighted city means for 1976 because of the combined weighting of adjustments for socio-economic and age-grade differences, the investigator concludes that there are sizeable differences in performance across the three phases of the study, with the highest scores achieved by the current group. The extent of the differences appears in Table 6:17.

Despite the fact that students in grades 5 to 7 today appear to outperform their counterparts of a generation or two ago in fundamental mathematics skills there is good reason to suggest that students still do not perform in some of these skill areas as well as might be expected. This conclusion is supported by an item analysis of scores on the Niagara Revision. For instance, one student in five in grade 5 missed questions on simple subtraction. When the subtraction questions dealt with fractions, two out of three students could not successfully perform the operations. At the grade 7 level only one-half of the students could successfully answer questions dealing with subtraction of fractions. At the same grade level two-thirds of the students missed questions using percentages, and one in three missed questions on simple addition with decimals. In a number of other subskills, including problem-solving, the results are even more negative, indicating that in spite of the apparent improvements in grades 5 to 7 arithmetic over the past forty years, there is still ample room for additional improvement.

On the basis of these findings, the investigator offers some reassurance to those who wonder whether present-day students perform as well in mathematics as they did in the past. At the same time, the disconcerting reality is that even with the relatively superior performance in grades 5 to 7, students still do not perform remarkably well in a variety of fundamental operations, and grade 8 students perform significantly worse than their counterparts in the 1930's and 1950's.

Some of the apparent reasons for the contrast between achievement in grades 5 to 7 and achievement in grade 8 have already been presented. The most influential change accounting for this apparent paradox has been the reduction in the emphasis on mathematics in grade 8 following the removal of the high school entrance examinations, with the result that mathematics today receives no greater overall emphasis in grade 8 than it does in the other grades under study. By examining achievement in arithmetic computation and reasoning only it may be that the study gave a less comprehensive look at mathematics in grade 8 today than did the more general test at the grade 5 to 7 level, and it may well be that achievement in the two grade 8 areas examined is relatively weaker than in some other dimensions emphasized in modern grade 8 mathematics programs.

These considerations, however, do not adequately account either for the reduced levels of performance in grade 8 or the still somewhat unsatisfactory levels in grades 5 to 7.

The major sets of concerns in the mathematics area revolve around questions of objectives and content of mathematics programs on the one hand, and issues related to pedagogy on the other.

A central issue is whether mathematics in general and the particular skills in question in this study continue to represent as important a dimension of schooling as they did formerly. The addition of more subjects to the school program, the reduction of time allocations to mathematics programs, and the large amount of new content in modern mathematics curricula suggest that the skills under study find less importance in schools today than heretofore. In view of the fact that the Ministry of Education and the local Board of Education have mandated the teaching of the skills under study, and since students do less well than expected, educators should take a new look at performance in these areas, to consider not only the relative importance of the objectives themselves, but also the levels of performance that should be expected. Modern mathematics programs may quite properly contain a variety of competing content that at times takes precedence over fundamental operations. However, in view of the sequential nature of mathematics skills, the wisdom of neglecting the attainment of levels of mastery in fundamental operations higher than those documented in this study is open to question.

There is a strong possibility that the profession is responding to the wrong concerns resulting from public criticisms about the need to get "back to the basics." This study offers an excellent data base for inferences about relative gains in performance in certain basic skills areas. The fault that may exist in the response to public sentiment on the question is that educators may be responding only to the feeling that students in the past may have performed better in these skills. The investigator suggests that educators look at a much larger issue, namely: should students be doing better than their present performance indicates? There exists a strong possibility that public criticism is not about the comparative abilities of students today and yesterday, but about the concern that students are not performing up to the levels of expectation that seem reasonable today. That is, public criticism may be directed both toward the objectives of mathematics programs and the standards of performance that are achieved and apparently accepted in schools. This interpretation of public criticism suggests that what the public is really saying is that a greater proportion of the mathematics program should be focused on fundamental operations, and that the expected levels or standards of performance should be not merely as high but higher than they were a decade or two ago.

Moreover, educators may need to review the implementation of Ministry and Board policy to determine the grade levels at which satisfactory standards of performance appear to be breaking down. An examination of documents from these two bodies satisfies the investigator's inference that the basic skills in mathematics are not only included in curriculum outlines, but also are given a prominent position. If in practice, as our data suggest, students are not doing as well in the fundamental operations as the documents imply, a careful examination of achievement across the grades would appear to be worthwhile and perhaps essential. Thus, public expectations, as expressed in Ministry and Board guidelines, and accurate information about student performance are essential in revising plans with respect to the balance of content in mathematics programs, the pedagogical emphasis that should be placed on each component, and the acceptable levels of performance.

One area of mathematics receiving considerable emphasis today is problem-solving. Considering the attention paid to this skill, one might expect that students today would outperform students from the earlier phases of the study by a wide margin. In the Niagara Mathematics Test Series, Arithmetic Computation and Reasoning, items were included that were categorized as problem-solving. Yet even on these similar problem-solving items students in 1976 did not outperform students from a generation or two ago. The implication, already alluded to elsewhere in the report, is that a fundamental difference has developed with respect to the nature and purpose of problem-solving, from a fairly stereotyped procedure used formerly to solve specific types of problems to a more creative approach to dealing with problems for which the student has no standard algorithm, and wherein he/she must employ a novel or unique strategy. While the results show that the students in the earlier two phases of the study had higher levels of performance on the stereotyped problems of those decades, there is no evidence to suggest that they could also outperform students today on the types of problems now used which are more appropriate to real-life problem situations.

Turning from questions of program content to pedagogical issues, one must consider some of the implications of highly structured versus reasonably autonomous classroom practices. The training of students to write grade 8 entrance examinations had the effect, in the 1930's (with some continuation into the 1950's), of standardizing instruction. Students in the various schools of the study typically experienced the same textbook, a structured program of studies, and fairly carefully controlled classroom experiences. It was reasonable, then, to expect fairly high and fairly uniform patterns of achievement on those skills emphasized both in the program outlines and in practice. The Dominion Tests given to grade 8 in the 1930's and 1950's readily matched the uniform content emphasized in the schools. Other practices designed to get students ready for entrance examinations, such as longer sessions of mathematics, test-writing practice, drill on set question types, the study of previous entrance examinations, and the like, also had the effect of inflating grade 8 scores over those attained in the lower grades. The reduction of these forms of emphasis, changes in the supervision of teachers, the addition of new curriculum components in mathematics, and greater diversity and autonomy in pedagogical practices in classrooms can be expected to have produced greater diversity in student competencies, accompanied by a lower general standard of performance in the skills area in which emphasis and conformity were typical in the earlier periods. The implication of this argument is that the more diverse course content and pedagogical practices in mathematics today, while contributing to decreased proficiency in the two skills areas under study, also probably contributes to increased proficiency in areas today that were not a part of the mathematics curricula of the earlier periods. The contrast between the comparative achievement of students in grade 8 today and achievement levels in grades 5 to 7 appears to support this argument. In short, there is considered to be little or no need today to give greater emphasis to mathematics in grade 8 than in grades 5 to 7. As a result, the actual achievement levels in grade 8, while considerably below those of the earlier periods, appear to be fairly similar to levels in grades 5 to 7, which themselves are ahead of those of the earlier periods.

The data today suggest that there is a much greater range of achievement both within classes and from class to class and school to school than was the case with the more highly structured and conforming programs of former times. Without wishing to imply the need to return to highly structured and centralized teaching practices, the investigator nevertheless suggests that there would be some merit in achieving consensus across the system as to the major priorities in the mathematics programs at various levels, some agreement on acceptable levels of achievement of students with comparable abilities, and increased professional development focus on practices, materials, and organization that contribute to higher levels of student performance in mathematics.

While recognizing the contributions of modern pedagogical techniques and class management practices to the all-round development of students, the investigator wonders to what extent these more democratic, student-centred, and less rigidly controlled classroom conditions have contributed to the discouragement both of drill and of the use of frequent short testing sessions. While the present investigator supports practices which create happy and creative classroom climates and experiences, he has reason for concern about ways in which these intentions in some classrooms have competed with or almost obliterated sound pedagogical practices requiring teacher-centred or teacher-directed activities. Excessive attention to students' self-direction may have made teachers unwilling or even unable to introduce students to the most efficient pedagogical practices for certain kinds of performance. Drill, for instance, is often challenged because it conjures up images of teacher dictatorship and the lack of awareness of the individual needs of students. What may be needed are procedures for making drill more varied and interesting and classroom organizational patterns which enable teachers to adapt drill techniques to individual and small-group practices.

Another pedagogical issue has to do with the investigator's concern about socio-economic effects on learning. Several of the schools in the study represent the lower half of the socio-economic continuum, at least in relation to the rest of the city. Yet the investigator questions whether these students are given special consideration in terms of content relevance, adapted pedagogy, and additional supervisory support. Indeed it may be the case that classes in 1976 in these schools, since they are more heterogeneous in student achievement levels than in previous decades, require a wider variety of learning environments if students are to learn at rates comparable to those attained by their counterparts in the more homogeneous classes of the past. While recognizing the financial constraints that school boards are experiencing in the mid-1970's, one must consider whether proportionately greater or better resources must be allocated to schools in lower socio-economic situations, whether by class-size reduction, the deliberate recruiting or allocation of the best available teachers, or the addition of materials and supervisory help. Reference was made in Chapter 4 to the fact that these schools have had a proportionate reduction in supervisory support. Careful consideration and study should be given to the question of the kinds and amounts of supervisory assistance that contribute most directly to student gains. Then decisions should be made with respect to allocating supervisory assistance, including the direct attention of principals toward improving student performance, proportionate to the needs of students in various types of school environments.

Three other issues arise in analyzing the comparative data on student performance in mathematics across the past forty years in these schools. In focusing on comparative achievement data, the investigator and probably those who may wish to study, compare, and use these findings are naturally forced to take an essentially backward look at what happens in schools. Therefore, in studying these conclusions, the reader must maintain a balance between comparisons with yesterday and the obvious need to prepare students for the world in which they will have to apply the skills acquired at school. Too great a preoccupation with comparisons with yesterday obviously would shortchange the student today, unless one seriously believes that we should return to the harsh realities that accompanied the less complex times nostalgically referred to as "the good old days." A more rational approach to the problem of looking backward and forward is to deliberately identify and retain those traditional skills which continue to have meaning today and are likely to be fundamental in the years ahead and to ensure that they are achieved to a level which will provide for their efficient use, not only in independent and isolated contexts but also as dependent steps in a well established and clearly integrated continuum or "ladder" of skills. It is not difficult to sustain the argument that both computational skills and mathematics problem-solving skills will find a predominant place in such an assessment of objectives worthy of retaining at high levels of mastery. Indeed one could argue that both of these areas, and particularly problem-solving, should demand higher rather than constant standards of achievement today and in future. Today's student will be expected to apply mathematics both in the relatively simple context visualized by "corner store" mathematics objectives and in the increasingly complex and varied world of modern living wherein not only the uses but also the very definition of mathematics itself continue to broaden.

Much of the current concern of parents with respect to mathematics achievement is a reflection of a communication gap existing between schools and homes. The schools in this study now have data with which to reassure the parents concerning the comparative levels of achievement of their youngsters in school, and the investigator contends that the schools should make this information available to parents. Such information by itself will not, however, completely allay the communications problem. The sudden movement into the "new math" program a decade ago reduced and in many cases eliminated the involvement and intervention of parents in their children's mathematics studies and particularly their homework. Prior to the introduction of new math parents were comfortable with their children's mathematical practices, because they were similar to those that they remembered having themselves employed. Suddenly many components of the mathematics curriculum, such as set theory, became completely foreign to parents. The effects of the change on parent interventions and parental attitudes is worthy of more formal study, but can be expected to have contributed negatively to student achievement in the most recent phase of the current study. It appears unlikely that the problem created by the introduction of new math will be repeated. Parent-teacher interviews, the involvement of parents on curriculum committees, and the use of parents as volunteers in schools, are practices which will enable many parents to develop insights about new mathematics content and practices and may encourage them to renew their traditional intervention into areas of the mathematics curriculum with which their children need additional help.

Finally, educators concerned about the teaching of fundamental mathematical operations and the significance of these basic skills in the modern world will need to establish rational policies about the extent to which students may rely on pocket calculators to perform these functions. An untenable paradox is created in schools where teachers refuse to permit students to use such devices, when they themselves employ them in calculating students' marks. If students can do accurate operations with these devices, need the teachers spend the same amount of time as in the past on drilling fundamental skills? Is it possible that the time saved may have to be spent on greatly improving the understanding of the operations involved in order that students can make the most effective use of such technical devices? Is it reasonable to expect that the improvement in understanding processes and inventive approaches combining calculators and routine drill will enhance both the speed and accuracy of computation? There appears to be little reason to suggest that the more flexible use of calculators by students, even in elementary grades, will cause a deterioration in their basic skills operations, provided teachers program both skills with appropriate balance; nevertheless, the introduction of a sweeping innovation of this kind should be monitored to ensure that desirable understandings and mental skills in computation not be endangered.

ACHIEVEMENT IN LANGUAGE ARTS

In this section the conclusions presented in Chapter 7 are restated, along with a number of implications. For the sake of both consistency and clarity, results from the three different sets of tests are considered in the same order as in Chapter 7, namely, grade 8 and grade 6 reading, and then vocabulary in grades 5 to 8.

From the administration of the Dominion Group Achievement Test Paragraph Reading, Forms A and B, the study showed that grade 8 students in 1975 performed at a level slightly higher than their counterparts in the first two phases of this study. Although students in the present phase did slightly less well than their counterparts in 1952 on Form A, this loss was outweighed by their superior performance on Form B. In Chapter 7 reasons have been presented for the assumption that Form B is a superior test to Form A for the current phase of the study.

The finding that student performance on the two parallel forms of the test changed differentially over the past forty years introduces an unusual situation. Because the tests have changed unequally in validity over time, the investigator hypothesizes that it is the educational contexts that have resulted in the differences. Apparently a variable not predicted or planned for in the construction of the original tests has surfaced, thereby reducing the reliability of tests originally considered as parallel. Further study will be needed to detect, analyze, and account for this unanticipated factor.

Another aspect of the findings about reading at the grade 8 level is reflected in the investigator's concern about what the Dominion Tests actually measure. On analysis the tests appear to deal with items of simple memory or recall, then with basic comprehension, and later in the tests with higher levels of comprehension. In at least the last few items of the tests the student is required to deal with levels of inference far beyond the usual ability levels of grade 8. In spite of these concerns, the tests do measure a number of skills important to reading proficiency at the grade 8 level.

One might look at the content validity of the tests from another perspective. What kind of reading constitutes the largest proportion of student reading material in the mid-1970's? Changing patterns in pedagogy and class management, greater student autonomy within the classroom, the replacement of one basic reading or literature book with a variety of books, and a greater array of optional reading materials are characteristics of the modern classroom which contrast with the pedagogical, content, and process components of reading in the 1930's and 1950's. While the investigator initially considered these tests to be reasonably suitable for purposes of comparison, a more critical examination suggests that there is reason to question the validity of the Dominion Tests of Paragraph Reading as tests of fair comparison at this time, on the grounds that the use of relatively uniform and classical patterns of expression focuses on only one style of reading, and one that is atypical of the common reading experiences of students in the 1970's.

Furthermore, a number of components of a modern reading program are not touched by the tests; for example, many schools in the region, including some in the study, are teaching problem-solving skills through the reading program. By using models for dealing with problems in logic and decision-making, schools are introducing students to higher levels of reading comprehension. In many cases these skills are further developed in a program of critical reading. These aspects of reading were not dealt with in the Dominion Tests since they were not objectives in the reading curricula of the 1930's.

Considering the possible weaknesses in the parallelism and reliability of the two forms of the Dominion Tests given in 1975, one might conclude that the evidence that grade 8 students today perform about as well as, or marginally better than, their counterparts of a generation or two ago, based on performance on the same tests throughout all phases of the study, speaks well for reading instruction today. Of course, the study does not demonstrate the competencies of grade 8 students in the wider array of reading skills. Consequently there would be value in preparing tests to describe student performance in these other areas, in order that future longitudinal studies can take into account this somewhat broader spectrum of reading skills.

Turning our attention to reading skills in grade 6, the study showed that on the average, grade 6 students today outperform in a group of reading skills similar groups of students during the three earlier applications of the Thorndike-McCall Reading Scale in 1938, 1953, and 1971. Similarly students in 1971 outperformed their counterparts in the 1930's and 1950's, and the graph (Figure 7:1) describing student performance on this test over the forty years of the study shows a steady but small improvement across this time span, represented on the graph by an almost straight line with a modest upward trend. Even after adjusting the mean scores downward in both 1971 and 1975 to account for effects of changes in age-grade levels, students in the 1970's remained at about the same levels of performance as in the earlier phases and in 1975 achieved slightly above any previous levels. Therefore, the investigator concludes that at the grade 6 level students today outperform, across a range of reading skills, their counterparts in all three earlier phases. Because of the particular qualities of this test and the fact that it was administered without revision in all four phases, the results are of particular interest. For a number of reasons it could be argued that the pattern represented by scores on the grade 6 reading test is more representative of comparative achievement levels in language arts across the years than that for any of the other tests.

Despite some invalid items in this test, it was considered suitable for use in all three phases of the study. Also it appears to test a much wider range of reading skills than the Dominion Group Achievement Tests used at the grade 8 level. One interesting anomaly in the use of this test is that there is only a negligible socio-economic difference between performance of students in the various schools of the study. A number of tentative hypotheses that might account for the apparent absence of socio-economic effects are: the contexts tapped by the test could be so limiting that the greater flexibility of higher socio-economic status students is not assessed by the items; there is a possibility that the lower socio-economic group, for reasons not determined by the present investigator, has a greater affinity for the contexts presented; the test content, apart from the context dimension, could differentially favour the lower socio-economic group; or a variety of current internal invalidities impossible to anticipate may be interacting, resulting in inconclusive data relative to the socio-economic factor. The data on these hypotheses are insufficient, making a definitive choice among them impossible.

An additional limitation of this test with respect to modern reading programs is its complete absence of material that would be considered Canadian content. The test items appear to be somewhat classical in style and seem to refer to English and American rather than Canadian contexts. In view of the use of Canadian content material in classrooms today, one might conclude that students in 1975 were penalized slightly by this factor.

Another area not adequately tested by the Thorndike-McCall Test is the set of skills typically grouped under the heading "critical reading." Had the test measured some of the more complex skills of comprehension, students today probably would have further outperformed their peers from the first three phases of the study.

Another issue relating to student performance in reading has to do with whether students today read as much as they did in the past. This question is complicated by two factors - the increased availability of reading materials in libraries, homes, and schools, and the competition for students' leisure hours by television and the wide range of their extramural and social activities. It is inferred that students engage less in leasure reading today, but television may possibly enhance their vocabulary development and some other aspects of reading. The reader is left, then, with an interesting picture of student performance in reading. While the Thorndike-McCall does not test the full breadth of student reading today, and while students apparently are reading less, the conclusion is reached that students are performing better than in the past on a fair test of a range of reading skills given greater exposure in the 1930's and 1950's. The seeming inconsistency in the logic can best be understood by a consideration of the quality of reading instruction and resources today. Part of the improved quality of reading instruction relates to the gradual disappearance of subject barriers. For example, one speaks today of language arts, meaning not only reading but also the interrelated skills of listening, speaking, and writing. Skills learned in one area of study are sequentially reinforced in other subjects, providing for a greater potential for skill acquisition. Part of the improvement in student skills today must also be attributed to improvements in teacher preparation. Most teachers of language arts in the study schools today are well grounded in the use of instruments for the diagnosis of reading problems and have additional support from competent language arts consultants, special education teachers, the regional office of the Ministry of Education, and Niagara Centre, O.I.S.E. Therefore, while student performance in grade 6 paragraph reading, representing a variety of skills, is higher than in the previous periods, other improvements in the system imply that there is continued room and opportunity for further improvement in reading skills.

The third area of language arts under study was the ability of students to generate appropriate vocabulary. Students in grades 5 to 8 in 1975 performed worse than their counterparts on the original form of the Vocabulary Test, Number I, Ninety Common Names; however, the results from the administration of the Niagara Revision of the test demonstrated that students in these grades today have skill in generating vocabulary superior to that shown by students in the two previous phases of the study. The latter conclusion requires additional comment in light of the fact that the investigator has identified several related issues. A number of these issues have been discussed in appropriate places earlier in the report.

One of the fundamental concerns about the original vocabulary test relates to its suitability in the 1930's and 1950's. Both of the earlier researchers expressed doubts about the suitability of the test at the time, and these doubts were substantiated by an examination of the scores obtained by students and by the fact that adults today have considerable difficulty with some of the items on the original test. The degree of suitability of the test in the 1930's and 1950's is not precisely established. However, there is no doubt of its unsuitability in the 1970's, as evidenced by the disappearance of some items from child vocabulary, changes in the meanings of certain words that have persisted, and the present wider range of alternative answers to some of the items.

In recognition of the inappropriateness of the original test, the present investigator and a competent panel of educators constructed a new test for administration to grades 5 to 8 in 1976. Even after adjustment for socio-economic effects and age-grade differences, the modern students outperform their counterparts of earlier periods by a wide margin. In addition, students today are able to generate a much wider range of acceptable alternative answers.

Also, an interesting and perhaps significant area of investigation for future researchers will be the study of factors that account for some students today knowing items that are considered to be antiquated. What collection of common experiences might account for students in

1975 knowing a reasonable number of vocabulary items no longer employed in the culture? If this information were known, useful conclusions could be developed regarding acquisition of new words.

Moreover, the variety of acceptable alternative answers given by students is evidence that the distinction between formal and informal language usage may be narrowing, at least among school children. This tendency may result from the constant bombardment of informal language through television, accompanied by a shift away from leisure reading, which at one time focused on the classics.

It is logical to assume that early exposure to television enlarges one's vocabulary. This is undoubtedly the case in early childhood. Indeed, one frequently hears that most of a child's functional vocabulary is learned apart from and before entry to school. Therefore, while there is evidence in this study that children's vocabulary has improved and broadened, a share of the credit for the superior performance must be given to factors beyond the school. The in-school influences contributing to improved vocabulary development include a wide range of interesting textbooks, visual aids, and teaching methods, all of which can be expected to contribute to the more formal use of language.

In drawing the conclusion that students today outperform their counterparts from a generation or two ago in the initiation of vocabulary, one must point out a weakness in the study with respect to the quality of the student responses to test items. The marking scheme simply accepted or rejected words on the basis of whether they were correct responses, but failed to discriminate on the basis of the quality of the responses. While the student today exhibits increased quantitative use of vocabulary, the study presents no evidence of the comparative quality of the words selected, apart from whether they were correct or not.

One final observation may be somewhat disturbing to the reader. An examination of the tables in Chapter 7 illustrates that the greatest comparative improvement in student performance in vocabulary is at the grade 5 level. In subsequent grades, the amount of increase relative to the earlier phases of the study becomes less and less.

That is, while grade 8 students today outperform their counterparts in the 1930's and 1950's the margin of difference is less than at the grade 5 level. Does this pattern indicate that as students proceed into the secondary school, the margin of superiority will continue to decrease and ultimately vanish? No conclusion is presented concerning this question. However, should the possibility stand a test of empirical study, the investigator suggests that comments presented earlier concerning the breakdown of the barrier between formal and informal language, and the effects of television on language development of the young child, may contribute to the explanation. In the early years, improvements in vocabulary result primarily from greater and earlier exposure to a variety of language. In this regard, the effects of television were stressed. In higher grades, the increases in performance relate more and more to vocabulary normally outside the range of common usage, and therefore result mainly from school experience.

THE TREATMENT OF MAJOR VARIABLES

In addition to the conclusions concerning student performance in mathematics and language arts, a number of other conclusions and implications require further description and explanation. It should be reiterated that the design of this study, in attempting to account for changes in student performance, looked at four major variables - socio-economic factors, age-grade levels, vocabulary, and school objectives - and attempted to account for the effects that they have had on student achievement. It is the intention of the following section to look at each of the major variables in turn.

The investigator reacted to the research literature which appears to be of one accord regarding socio-economic levels and their differential relationship with student achievement. Indeed the relationship has such wide acceptance that it was obvious that the current study should be designed so that relative socio-economic levels could be used as an important variable accounting for changes in student achievement between the phases and between schools in the current phase. However, apart from there being general agreement that socio-economic factors affect student achievement differently in low and high areas, no single research study forwarded a formula suited to the quantification of the socio-economic variable that could be used in this study. The closest the investigator came to a realistic and logically valid formula was a suggestion in Sewell and Houser (1975): *Education, Occupation and Earnings: Achievement in the Early Career* (pp. 49-50), that there is in existence an "index of socio-economic status that is a linear combination of parents' average income and educational attainments with fathers' occupational status." For a number of reasons, such an index would be impossible to obtain for the current study. To gather the data for such a large population would be an expensive and time-consuming task even if a mandate were given to enter the privacy of family lives to investigate incomes, educational attainments, and occupational status.

Thus, as described in Chapter 4 it was decided that socio-economic data would be obtained from general descriptions of a cross-section of informed persons in the community; then, based on these data, socio-economic adjustments to raw score means would be applied to all student mean raw scores for all grades and all tests in 1975-76. In addition the same indexing was used, with required explanations, for the Hastings data from 1971 on grade 6 student performance in reading. An examination of Table 4:8 and the summary tables for each of the tests shows *that there were, with one exception, differences in achievement related to the relative socio-economic levels employed in the study, and that appropriate adjustment factors were therefore possible to calculate*.

The investigator recognizes that while these adjustments may be logically valid and have some statistical justification, they remain only a rough approximation of relative socio-economic effects on student achievement. The rational validity for the indexed adjustments is apparent; the differences were calculated from mean raw scores of schools identified as being below average, average, or above average with respect to relative changes in socio-economic status. Had the study been able to afford to gather a wider set of demographic data on the homes of students, a more sophisticated adjustment formula probably would have been developed. Considering the general agreement in the literature on the influence of socio-economic factors on achievement, and the evidence that relative socio-economic levels change in some instances over periods of one or two generations, it can be expected that even greater attention to this variable will be given in future longitudinal studies, and that the researchers involved will develop socially acceptable and empirically valid measures to discern with even greater accuracy the effects that socio-economic changes have on comparative student performance.

The second major variable introduced as an adjustment in the present study is the age-grade factor. An important finding of this study, reported in Chapter 4, and developed as a result of painstaking attention to student records, is that students today in the grades under study are older than their counterparts in the 1930's and 1950's, and also are slightly older than would be determined on a model developed by assuming uniform promotion of exactly one grade per year. This finding indicates that the amount of retardation in the schools under study is greater than the amount of acceleration, a fact which appears to run counter to the popular rhetoric about promoting students in keeping with their readiness. Because the average age-grade of students today in the study schools is higher than in either of the earlier phases of the study, obviously the adjustment factor in each grade has the effect of reducing the scores in the current phase. Had the investigator wished to introduce an inconspicuous bias into the results concerning students today, it would have been relatively easy to avoid the consideration of this variable. To the unsophisticated reader, it might seem encouraging to read that grade 7 students, for example, outperform their counterparts of a generation or two ago on several specific tests or in certain skills. However, such a conclusion would have been irresponsible in the light of the finding that the groups are different to begin with in respect to age, if this variable had not been treated. Thus, a fair comparison of student achievement demands that statistical adjustments be made to account for the variation in age-grade. The result of the adjustment, therefore, is that comparisons are made of students of the same age, registered in the same grade. The rationale for this adjustment is obvious, in view of the fact that when one wishes to make comparisons over time a similarity in age is implied.

The system for quantifying age-grade differences is based on the actual differences in the ages of students writing the test. By simple computation, the advantages in test scores attributable to the greater age of the modern students could be subtracted from mean raw scores. In one respect at least, this procedure is more defensible than that created for socio-economic indexing. For socio-economic differences the investigator had to rely on highly subjective judgments about relative socio-economic levels of the homes of students. Even with high inter-rater consistency, the use of only three categories coupled with the personal inferences of judges is open to some question. With the age-grade differences, however, no such inference was required. The ages of students in each of the grades is fact! The only possible criticism of validity would relate to how age-grade adjustments were calculated. With respect to this latter possibility, the investigator has been unable to find a better way of indexing differences in age-grade than through a computation that employs actual age differences and student performance on tests written.

The third variable, vocabulary, unlike the preceeding concerns, poses unique problems for anyone involved in a longitudinal study spanning two or more generations. While it is well known that child language changes, and that the meanings of words are altered, there is little systematic documentation of these changes useful in assessing tests created formerly. Even with careful documentation of such changes, it would be a complex task to assess fully how the wording on test items themselves and on the accepted answers would be affected by these changes. For instance, how should the investigator make judgments about the difficulty level of words used in the first two phases of the study which have disappeared from the modern idiom? Then, having decided that words such as "scuttle", "bodkin", and "salt cellar" have all but disappeared from child vocabulary today and therefore need to be removed from a current test of fair comparison, what words of relative current use and difficulty should replace them? In attempting to assess the effects of changes in vocabulary the investigator of this study was faced with other unique problems. One of these was the fact that the original vocabulary test of 90 Common Names was never standardized. Also, it was developed in England, and therefore was unable to take into consideration, even in the first instance, particular aspects of Canadian vocabulary. The hypothesis that students today have broader and more varied vocabularies as a result of increased exposure to the media, class discussions, and reading materials could not be tested with the instrument used in the earlier phases, because of both the limitations of the test itself and the narrow range of acceptable answers in the original teachers' manual. Unanswered concerns about the validity of the instrument during the first two phases resulted in questionable baseline data for use as points of reference for the current study.

In the light of these concerns, the original test was administered in 1975 in order to check the investigator's judgment of its unsuitability. Then, a necessary replacement test conforming to item types of the original test and to the modern idiom was constructed. The comparative results of student vocabulary achievement in grades 5 to 8 for the three phases of the study show students in 1976 clearly outperformed their counterparts a generation or two ago. The extent of these differences, with respect to both the number of correct answers and the range of alternative acceptable answers (described in detail in Chapter 7 and illustrated in Tables 7:8 to 7:9(d)), are sufficiently great that they were accompanied by a set of explanatory notes.

The extent of the importance of the vocabulary variable, as evidenced by changes in performance on the vocabulary test, is readily apparent when the test concerns vocabulary only. However, the effects of this variable do not appear only in tests of vocabulary. In the Dominion Group Achievement Test of Paragraph

Reading, and in both the items and the marking key for the Thorndike-McCall Test, the vocabulary used sometimes represented the vanishing idiom of a previous generation. While the specific effects of such difficulties in language are not accounted for in this study, and while no adjustment in scores was made for them, the reader is encouraged to recognize their existence and infer the effects they may have had on students' scores. It seems reasonable to propose that had all the items, instructions, and scoring keys been in commonly used modern language, the scores in the current phase, particularly in the language arts skills, would have been even higher.

The final variable, school objectives, probably affects the data on student achievement to a greater extent than any other single factor. Various factors directly influenced by changes in school objectives, such as course content, emphasis on subjects or components of subjects, amounts of time devoted to each area, etc., all can be expected to have their own impact on achievement. The attempt to treat all of the factors stemming from changes in school objectives, and to weigh their differential effects on student performance, becomes an almost unmanageable task. A simple example will illustrate the difficulty of the problem. This study does not attempt to equate the specific amounts of time spent by students in the three phases of the study on the particular language arts and mathematics skills measured by the tests. Therefore, if a student today does slightly less well than his counterparts in 1934 and 1953 in arithmetic computation, but spends considerably less time in arithmetic drill, what conclusion is to be drawn regarding student performance?

Longitudinal studies based on tests used in the earlier phases and geared only to student performance on topics common to all phases of the study obviously favour the results from the earlier phases. Ordinarily tests are constructed to reflect curriculum emphasis at the time of their use. Changes in curriculum content typically give different weighting to different topics at various periods of time. Thus, as the curriculum content expands in tune with the complexity of life, the investigator is left with longitudinal comparisons based solely on the curriculum content of previous generations of students. In the language arts area, the study did not compare critical reading, creative writing, discussion skills, or research study skills because these areas, now emphasized in curriculum, were considered less important a generation ago and did not constitute parts of the tests administered at that time. Looking ahead, the investigator imagines a time in the future when some of the skills examined in this study may no longer be of central interest to educators. In order to cope with this variable, future investigators are encouraged to make realistic comparisons not only across the various phases of the study but also across content areas, making sure to include content recently introduced into the curriculum. In some cases, these later contributions may eventually become the essential elements for subsequent comparison. The current study, in reconstructing original and creating replacement tests, attempted to apply this principle by offering to future investigators data in two areas of current concern - problem-solving in the Niagara Test of Fundamental Operations, and current vocabulary in the Niagara Test of 90 Common Names.

EFFECTS OF OTHER VARIABLES

The foregoing provides the general findings of the study and also accounts for the influence of four major variables - socio-economic factors, age-grade levels, vocabulary, and objectives - on comparative student performance. Although the four variables that have been treated are considered to have had the greatest impact on student performance, and therefore to account most for changes in student performance particularly during the past twenty years, some attention must be given to a wide range of other variables for which changes in school systems over the past several decades can be readily documented, and which can be expected to have had effects, either positive or negative, on student performance. It is much easier to identify these changes than it is to account precisely for the effects they have had on student achievement. Because no systematic treatment or measurement of these variables was possible within the design of the study, the stated direction of the effects themselves and certainly the suggested degrees of effect, are essentially hypotheses, based on the researcher's general experience in schools over the past three decades, and related as far as possible to observations in this and other evaluation studies.

Although many of these variables have complex inter-relationships with others on the list, an attempt is made in this section to isolate them for purposes of identity and to indicate directions of their effects, and in some instances to estimate the dimensions of their influence.

1. Class Size. The amount of attention given by teacher groups, particulary during salary negotiations in recent years, to teacher-student ratio, or class size, indicates that many teachers believe differences in class size almost automatically produce differences in average student achievement. The research literature in recent years is rich in research studies on this question.

Many of these studies indicate that the popular notion that a reduced class size produces improvements in average student gain is generally not valid unless other important changes in the teacher variable take place. Average class size in the school system under study dropped during the forty years of the study, from 40.4 in the 1930's, to 36.0 in the 1950's, and to 27.9 in the mid-1970's. However, the lack of documented information on changes in individual teachers' practices, coupled with the very heavy rate of change-over in teaching personnel, prevents the researcher from obtaining the data that would be required to conclude that reduced class size of itself produced marked changes in teaching practice or in student performance. Therefore, the hypothesis that a reduction in class size accounts for some of the gains in achievement, while logically defensible, cannot be scientifically supported from the data in this study.

2. Teacher Education. Few if any changes in schools in the past decade have been as important as those in teacher education. Improvements in this factor have almost certainly brought about concomitant changes in classroom methodology, and ultimately in student achievement, and can be expected to continue to have these effects for some time to come. The most important discontinuity in the long and often interrupted upward trend in teacher preparation was the requirement in the early 1970's that all teachers must have at least a bachelor's degree prior to entering the profession. Because almost all the students tested in the current phase were already in school, and some of them were half-way through elementary school prior to the introduction of this improvement in teacher education, it can be expected that the effects on student achievement have not as yet developed to the full extent that should be apparent when the majority of teachers have achieved this standard. While the researcher is prepared to attribute some of the improvement in student performance to improvements in teacher education, one must also qualify this generalization in the light of the two subject areas under consideration. For several generations, teachers have regarded language arts and mathematics as the basics of the school program, and therefore have had both the commitment and traditional methods for teaching these subjects with relative efficiency. Therefore, changes in teacher education have no doubt produced relatively greater improvements in certain other areas of the curriculum, in which traditional teaching practices have not been as well developed, and where in general there has been less commitment. It could be argued that the greatest improvement in the application of better teacher education has been in those areas of school life that relate to the all-round development of the child, such as discipline patterns, individualization, grouping techniques, and subjects such as physical education, art, music, science and social studies, and that there has been less improvement in the two more traditional skill areas.

3. Teacher Autonomy. Improvements in teacher education, enhanced professional security and status, the effects of Ministry guidelines related to the role of the individual teacher, and the recommendations of reports such as Living and Learning have created far greater possibilities for and actual measures of teacher autonomy and individual teacher control over curriculum than existed in the past. Logically, it would appear that these greater measures of autonomy would produce changes in teachers which would flow in both directions - on the one hand toward greater pupil achievement, and on the other away from it. Also, it seems logical to suggest that teacher autonomy would be exercised most strongly in new and creative subject areas, rather than in basic subjects such as language arts and particularly mathematics, where traditional teaching practices and considerable control of content have prohibited the exercise of large measures of autonomy. Therefore, we conclude that changes in teacher autonomy probably account for relatively little of the difference in average student performance in the subject areas under study, and that such changes occur in both directions and probably balance each other.

4. Classroom Methodology. It is impossible, for purposes of analysis, to divorce changes in classroom methodology from a wide variety of interrelated factors, including teacher preparation, materials, class size, grouping patterns, and even educational philosophy. As with teacher autonomy, it can logically be expected that the greatest effects of improvements in classroom methodology would be reflected in methods and subject areas where traditional practice has been neither well formulated nor socially controlled. Most observers of the school scene probably would agree that there have been important changes and improvements in disciplinary practices in schools, in the use of media and direct experiences, and in instruction in art, music, physical education and science. It would be less easy to gain such agreement on changes and improvements in methodology in mathematics and language arts. Nevertheless, there have been important changes, many of

which could logically be considered to be improvements, in instruction in these two areas during the past few decades, and it is reasonable to infer that these of themselves have produced positive, but probably small, overall improvements in student performance.

5. Grouping Patterns. Because grouping patterns are employed more widely and consistently in mathematics and language arts programs than in other phases of the curriculum, it can be expected that the achievement benefits that relate to such procedures will be most evident in these subjects. By and large, it can be anticipated that the use of grouping patterns related to students' achievement and ability levels will enhance achievement, and therefore must be considered positive factors in this analysis of other variables that may have influenced student achievement. However, the use of grouping may also have some negative effects on overall test scores, in view of the fact that it accommodates a wider range of students in each grade than would have been the case in earlier phases of the study. By making it possible to accommodate slower-learning students more readily than in former times, grouping practices undoubtedly account for many of the test scores in the lower range, and thereby have a depressing effect on the mean scores. In support of this point the reader should recall the evidence of the greater heterogeneity of classes in the current phase of the study.

6. Remedial Teaching and Special Education. As with grouping patterns, the focus of remedial teaching and special education is more concentrated in mathematics and language arts programs than in most other parts of the curriculum. Therefore the benefits of these programs can logically be expected to influence test scores positively. However, the provision of these practices has two other effects, desirable for general purposes but with negative effects on average test scores. These programs enable schools to promote students who in former periods would have remained at

lower grade levels, and also retain older and academically less able students who would have been "drop-outs" in earlier periods.

7. Teaching Materials. Improved teaching materials in greater number and variety must logically be expected to have positive effects on general achievement levels. Obviously the potential of improved teaching materials for enhancing student achievement is interrelated with a number of other influences, and without a design that would isolate the materials from associated factors, this study is incapable of measuring their effects.

One might propose, for example, that the addition of more colourful and interesting materials merely helps the school keep abreast of highly motivating counterparts in the culture at large, such as games, television, etc., which compete for the attention of the learner and create levels of expectations that the school must match if it is to maintain student interest.

8. Experience in Test Writing. In Chapter 5 a comparison is made of the type and amount of previous test experience, particularly with standardized tests, experienced by students in the three phases of the study. The reduction of standardized tests, the gradual disappearance of teaching practices designed solely to enhance the students' test-writing abilities, and the reduction in emphasis on formal testing in the 1970's probably have had no negative effect on student achievement *per se*, but can logically be expected to have had a negative influence on test scores.

9. Time Allocation to Subjects. Although no precise data on changes in time allocation have been gathered in the 1970's, in order to compare this factor with circumstances in the 1950's and 1930's, a number of changes in school programs indicate that there has been a reduction in the formal amount of time allocated

to mathematics as well as to language arts. The shortening of the school day, the introduction of rotary timetables, the addition of new subjects and new programs, and the increase in the number of professional development days are four changes which have put increasing pressure on the amounts of time given to these two subjects. A gradual shift toward environmental subjects and the creative arts have increased the proportion of time on these subjects, to the relative detriment of language arts and mathematics. It is logical to suggest that the reduction in time allocations to these subjects has had substantial negative effects. Probably the most dramatic evidence of effects of reducing time on any component of a subject area can be seen as a result of reductions in the periods devoted to drill on the speed and accuracy of simple number operations.

10. Increased Responsibilities of Schools. A further proportionate reduction in the amount of time allocated to mathematics and language arts can be attributed to the acceptance by the school of a number of additional responsibilities, including curriculum planning, diagnosis of learning difficulties, parent-teacher interviews, various kinds of referral, etc. These additional, albeit worthwhile, commitments of the school can logically be expected to have reduced instructional time and thereby to have produced a negative influence on student achievement.

11. Supervision. In Chapter 4, certain changes in supervision following the establishment of the county board were outlined. In the case of the schools in this study, the amount of supervision by central consultants, supervisors, and administrators was greatly reduced. Reference was made earlier in this chapter to the need to examine the impact of this change. Assuming that the assistance of supervisors has direct or indirect benefits to instruction, and therefore to student performance, the withdrawal of these services can logically be expected to have produced negative effects on student performance. To some degree this reduction has been

compensated for in the partial reallocation of activities of other personnel, including principals.

12. Student Motivation. Although it is impossible to study motivation for comparative purposes here, there is some evidence that students are less motivated than previously with respect to performance in mathematics and language arts. Competition from television and other media and the frequent perception by students of these two subjects as not being glamour areas of the curriculum may have contributed to an apparent decline in motivational levels. Social promotion, despite its benefits, may have reduced some students' motivation to strive to avoid failure. The apparent lack of motivation of many of the students with respect to writing the tests themselves has been discussed elsewhere in this report. This behaviour, linked with lower general motivational levels, can be expected to have accounted for lower test scores and lower general performance than would be anticipated if the students were highly motivated both in the subject areas and in the writing of the tests themselves.

13. Homework. In his study Partlow referred to the reduction of homework in the 1950's when compared to the 1930's. Since the 1950's, and in keeping with the philosophy and recommendations in Ministry guidelines, there has continued to be a change in both the amount and the type of homework. It is no longer customary to assign arithmetic problems or drill activities as homework; this tendency, admirable as it may be for other reasons, can be expected to have had a negative effect on test scores, largely because of the reduction in total practice time. This reduction can be expected to show up most in such things as the rapid and accurate use of simple number facts and routine calculations in mathematics problems. On the other hand, many of the homework assignments in use today focus on library research and the preparation of projects. Because both of these types of homework increase the amount and significance of reading, it can be anticipated that this change would have a positive effect both on vocabulary and on reading performance.

14. *Television and Other Media*. A number of recent studies indicate that students in the grades tested actually spend more hours weekly in front of a television set than in a classroom. This experience, coupled with students' experiences with other media today, is considered as having an important influence on student motivation and performance; a number of studies elsewhere have documented some of these effects. By reducing practice time, long periods of attention to media compete with factors that enhance achievement; this negative effect would be expected to be most evident in dimensions of mathematics that demand periodic sessions of drill. On the other hand the greater exposure to media no doubt can be expected to have positive influences on creative expression and on the breadth of a student's vocabulary. This exposure may account more than school programs themselves for the much greater range of correct student responses on vocabulary items.

15. *Philosophy and General Practice*. Although difficult to assess accurately or in isolation from other variables, among the most sweeping and pervasive of the factors which have influenced student performance, both negatively and positively, in the mathematics and language arts areas have been the changes in educational philosophy and the classroom practices that develop from such changes. The elimination of the threat of punishment and failure, provisions for identifying and meeting student needs, the shift from content to process as the central emphasis in instruction, attention to the physical, social, and emotional development of the student in addition to his intellectual development, the friendlier atmosphere of the classroom, the recognition and accommodation of a wider range of student abilities, interests and achievement, and attempts to place the child in the centre of the learning scene, have all had positive effects on school life in general. However, some of these tendencies have shifted emphasis away from some elements of mathematics and language arts programs. While it could be argued that schooling generally is enhanced by such changes, they can hardly expect to have produced higher scores on tests which were constructed to reflect the philosophy and practices of an earlier period.

In summarizing the apparent effects of a wide range of other variables on student performance, one must recognize first of all the complexity and the interrelatedness of these and other variables not included in the analysis. Having done this, and having assessed the probable effects of many of these changes, there appears to be, in general, a rough overall balance between the positive and negative effects of these changes on student performance. One must bear in mind in such an analysis that student performance on standardized tests, and particularly tests of another era, does not describe the total pattern of student learning in schools adequately. Therefore, results from a limited number of tests, often outdated, applied to a few elements of two subject areas, should not be used as a basis for either approval or criticism of practices and changes introduced with the intention of improving overall student learning across a range of subjects.

RECOMMENDATIONS FOR FUTURE STUDIES

One obvious advantage of the present study is the comprehensive data base it provides for future researchers. From these data several significant studies could be launched.

For example, researchers from the Lincoln County Board of Education may wish to consider step-wise regression analyses that will enable them to make better predictions of student performance. Alternatively one might want to use these data in order to reduce the number of tests required and the size of the test population, and at the same time to give board officials reasonable overall indices of student performance in language arts and mathematics. In each of these instances, the prediction, monitoring, discrimination, and control of student behaviour could be improved by using and enlarging the existing data base.

Test norming could also be completed for those tests used in the present study that are best suited to local board objectives. The large bank of data already recorded and tabulated, gathered from a large number of students tested in highly controlled administrative procedures, lends itself to this possibility.

Similarly, with a few minor adjustments made to test construction for the third phase of the study, the Niagara Test Series could provide excellent tests for the local board. With additional application to schools in a greater variety of socio-economic situations and urban-rural settings, and school systems outside the jurisdiction under study, the tests could become standardized instruments suitable for use throughout the province.

Since several of the senior elementary schools in this study have open-space classrooms, a limited comparison could be made of student performance for students in open classrooms and those in traditional settings. While recognizing that there are too few open classrooms in the study schools for a broad generalization to be made, the existing data could serve as a baseline for pilot testing to enable a research team to decide whether a more complete research design warrants the additional effort.

Both the original tests and the tabulation sheets provide opportunities for making comparative judgments about the performance in each of the areas under study for boys and girls. Using these data, a study of the comparative performance of boys and girls in certain skills such as arithmetic computation, arithmetic reasoning, vocabulary and paragraph reading would suggest interesting and useful projects, such as those often used as qualifying research projects by graduate students.

The graphs presented to show student performance in grades 5 to 7 on the Niagara Test of Fundamental Operations suggest another interesting area of study. A closer match of the entries on the graph with content introduced at each grade level would give a reasonably clear picture of the efficiency with which new items are learned at each grade level.

Future researchers may be interested in determining the amount of students incidental learning. By investigating course content and test item content across all phases of the study those items dealing with content beyond that identified in school programs could be located. Students' performance on these items then could be looked at as a measure of extra-curricular or incidental learning.

The study poses both a challenge and a set of insights to researchers involved in future longitudinal evaluations. Throughout this report, the investigator has introduced issues of and responses to limitations in longitudinal evaluations. Therefore, numerous suggestions are made at appropriate points thoughout this report regarding such diverse elements of the task as the overall management of a large-scale project, methods of accounting statistically for socio-economic and age-grade differences, procedures for developing criteria for selecting and creating tests of fair comparison, suggestions for matrix sampling, strategies for validating test-marking procedures, and activities and designs for reducing test-learning effects and ensuring confidentiality.

SUMMARY OF CONCLUSIONS

1. In comparing student performance in mathematics and language arts in the 1930's, 1950's, and 1970's in a group of schools in the city of St. Catharines, the investigator found that changes in four major variables - socio-economic levels, age-grade levels, vocabulary, and school program and objectives - all had important effects on students' scores, and required appropriate adjustments in order that fair comparisons could be made.

2. Consistent with the literature on the question, the study found that mean scores were lower on almost all tests for schools in lower socio-economic areas.

3. The study found that the average age-grade of students today in all grades studied is older than in the first two phases of the study, and also older than that suggested by a model based on the assumption of a consistent match between years and grade levels.

4. Changes in vocabulary made one test obsolete, and influenced the suitability of items on a number of other tests.

5. Changes in school program and objectives represented the most important variable in accounting for changes in student performance over the forty-year period; as a result, for a number of the tests, comparable instruments of fair comparison had to be constructed and administered.

6. The grade 8 students consistently performed in 1975 at a lower level in arithmetic computation and reasoning than their counterparts in the 1930's and 1950's, both on the original tests and on the Niagara Revision.

7. The students in grades 5 to 7 today perform slightly worse than their earlier counterparts on the test of mathematics fundamental operations employed in the first two phases of the study; however, on the revised Niagara Mathematics Test of Fundamental Operations, reflecting modern content and objectives, they outperformed their earlier counterparts at all grade levels. Despite this apparent gain in performance, there is still a very wide margin for improvement in many of the groups of skills taught in all three grades.

8. In paragraph reading, the grade 8 students today perform about as well as or marginally better than students in the same grade in the 1930's and 1950's.

9. On the Thorndike-McCall Reading Scale, considered to be the best of the original tests in the study and also the one most frequently administered over the forty-year span, grade 6 students today perform at a slightly higher level than those in the three earlier administrations, and the results from all four phases of the study show a steady but small upward trend in student competence in the skills tested by this instrument.

10. On the original test of students' ability to generate vocabulary from descriptions on ninety common items, students today in grades 5 to 8 perform less well than their counterparts in the 1930's and 1950's; however, the test employed was considered to be obsolete in some respects and unsuitable in certain other ways. On the Niagara Revision of this test, which uses modern vocabulary, students in all the grades tested in 1976 outperformed their earlier counterparts by a wide margin, showing not only superior ability with respect to vocabulary initiation, but also a much wider range of acceptable alternative answers.

11. The changes in student scores over the forty-year span of the longitudinal study cannot in every case be attributed to the four

major variables alone. In addition to the effects of the four variables intensively treated, some of the changes in student performance can be attributed to a set of other variables, including teacher education, supervision, teaching materials, television, motivation, changes in teacher-student ratio, etc. No attempt was made to examine these additional variables statistically.

12. Apparent similarities between the schools in this study and comparable schools in other jurisdictions may encourage the reader to generalize the conclusions beyond the population in this project. In this event, the present investigator cautions the reader to make such generalizations only after giving adequate consideration to the empirical, historical, and practical restraints inherent in such a decision process. At the very least, the reader, in making a wider generalization, should be expected to consider the apparent differences as well as the similarities in the systems or populations being compared. For this reason, while the researcher recognizes the value of using these data, sweeping generalizations were deliberately avoided in this report.

9 Other Outcomes of the Study

In the previous three chapters this report has dealt with the specific objectives of the project, namely to compare student achievement in language arts and mathematics in a number of schools in St. Catharines with their counterparts in the 1930's and 1950's, and to consider the impact of four variables in particular on changes in student performance during the latter part of this period.

Aside from the conclusions referred to above, a number of other major outcomes of this study merit attention. These concern the applicability of the generalizations in other respects and in other school systems, the utility of the design for future studies, insights concerning the administration of similar longitudinal studies, possible advantages to school systems of such studies, specific recommendations for conducting parallel studies in future, and a description of a data bank for such purposes.

1. Response to Current Criticisms. It seems unlikely that the findings of the current study will either broaden or allay current criticisms concerning student performance in schools. Hopefully the study will be valued in such discussions both for the data and conclusions provided, limited as their breadth of applicability may be, and for the suggestions the study makes toward an objective, rational approach to an area of discussion and criticism wherein bias and faulty memory unfortunately all too often are substituted for comparative data.

While the findings may have some general applicability to similar jurisdictions across the province, in view of the fact that most school systems have gone through approximately the same stages of development and are broadly comparable with respect to objectives of the elementary school, the study itself does not state or demonstrate that such a comparison can be made. Therefore the reader may wish to reflect on the degree to which the findings in the current study can be applied to other jurisdictions, and the ways in which the study could contribute to similar undertakings elsewhere. The care with which this study has been done, and the period of time over which the total study has developed, are expected to give the reader confidence concerning the procedures and findings for those schools and students in the study. It would not be a particularly difficult task for other school systems to administer one or more of the tests employed in this study to determine whether performance levels in their own jurisdictions are comparable to those in the schools under study. What still would be lacking in such a procedure would be the evidence of achievement levels during earlier periods in these jurisdictions.

One must consider, in the light of the findings of this study, what are the underlying causes for current criticisms of student performance. One possibility may be that the perceptions of adults concerning levels of performance when they attended elementary school are inaccurate or distorted; the data from the earlier two phases of the study indicate that the achievement levels were not as great as adults today seem to imply in their comparisons of achievement levels today and yesterday. A second possibility may be that adults today believe that achievement levels in the various recent decades have been roughly the same, but have higher levels of expectation for students today in view of improved teaching practices, learning materials, public support, etc. In view of the fact that this study suggests that when certain variables are taken into account students perform in some subjects as well or better than they did in previous periods, and when one considers that the major variable in this analysis is school programs and objectives, one is forced to consider the possibility that what adults are criticising is not student performance levels per se, but rather the objectives which underlie present school activity What this suggests is that some adults may be proposing that we return to the objectives of schools of yesterday rather than concentrate on higher achievement of modern objectives. The frequency of criticisms concerning mathematical computation, grammar, and spelling tends to lend credibility to the notion that many critics of elementary schools today are more concerned about the school's objectives than its achievement levels.

2. Implications for Program. It is becoming widely recognized by the profession today that the major goals of evaluation include comparison not only for the purpose of accounting to parents and the public, and for possible remedial instruction, but also for the identification of components of the curriculum which may need revision or other attention. For this reason more educators will probably concern themselves with the curriculum implications of this study than with the comparative data. While the performance scores indicate commendable trends with respect to some programs of the school, they also point to areas where there is some weakness. From these generalizations curriculum leaders, whether at the school or system level (assuming there is some evidence that these generalizations also apply to their own situation), may wish to identify those elements of program which should be retained with their present form and emphasis, and other components which need strengthening. Underlying such decisions, however, are questions concerning the desirability of the objectives themselves. Where an objective is no longer important or useful to the life of a student, evidence that the student is performing poorly on this objective does not represent an argument for revision.

A reasonable approach to setting curriculum priorities starts by identifying the important objectives of the school, subject, or grade in question. Then procedures are applied to determine first the desired level of performance on each objective, and then the actual level of performance by students. This procedure permits the curriculum developer to measure and describe the discrepancy; based on knowledge of the discrepancy he or she can decide the priority for action to improve this part of the curriculum. The present study can shed little light on what the important objectives are or what the desirable level of performance should be. Rather, it provides on a very limited scale information about student performance with respect to certain objectives, and therefore provides the curriculum developer with the data on which decisions about discrepancy and subsequent action can be made. Thus the reader should not necessarily assume from the generalizations in this study that more attention should be given to such curriculum components as drill, computational skills, or problem-solving. Rather, assuming there is a match between the jurisdictions under consideration, the reader should decide how much importance should be placed on these objectives and what the desirable level of performance should be, and then determine how great a discrepancy is represented by the findings in the study. Relating this viewpoint back to public criticism of student achievement in schools, the study suggests that greater attention may have to be given to developing a consensus on the desirable objectives of programs in mathematics and language arts before decisions can be made to alter programs in the light of the findings of the study.

3. Local School System Interests. Although this study was contracted by the Ministry of Education, its greatest value will be to educators in the system in which it was conducted, namely the Lincoln Board of Education. The data and conclusions of the study will shed light on a number of questions concerning policy and procedures in the system, and no doubt will cause consideration to be given to a number of other questions.

A number of the schools in the study are in the heart of St. Catharines. The study provides evidence on achievement levels for these schools which may help shape policies concerning support, teacher-student ratio, supervision, special services, promotion practices, and other factors for which comparisons and decisions must be made with consideration given to schools in various types of communities.

In Chapter 4 reference was made to a number of administrative changes following the development of the county system. This study may provide information concerning possible effects of some of these changes. For example, one wonders about the negative effects of a large reduction in the supervisory services available to the schools in question as a result of spreading the limited supervisory strength across the enlarged system in order to ensure equality of opportunity. This raises further the important question concerning the quantity, frequency, and type of supervisory service which can best meet the need to improve student performance in schools. Although it is suggested in this study that the reduction of supervisory and consultative services probably has had a negative effect on student performance, the study does not establish clearly this cause-and-effect relationship.'

The school system in question administered, at the grade 4 and grade 7 levels during the past year, a comprehensive test of basic skills based on country-wide norms. The performance of students in Lincoln County was considerably above the norms at both grade levels. This evidence, compared with our findings concerning student performance in the two basic subject areas over the past forty years, should enable curriculum leaders in the system to compare or verify information about general achievement levels in the system and to make appropriate modifications in objectives or emphases. Local educators, aware of other policies, practices, and trends in the system, will find other points of interest in this report. For example, the evidence on age-grade averages and acceleration and retardation practices will cause some reassessment of current uses of acceleration and enrichment.

4. Age-Grade Levels. The findings of the study in the area referred to above, namely the increase in age-grade averages and the current balance in promotion practices away from acceleration, raise questions about the rhetoric of continuous progress. The evidence suggests that very little acceleration occurs at all, with almost none beyond the fourth grade. The benefits of enrichment for large numbers of students apparently has led to a fairly uniform pattern of social promotion. Without launching in this report a discussion of the respective merits of acceleration vs enrichment, the author suggests that some schools may need to look at their balance between these two procedures for helping students grow at a rate reflecting their abilities, interests, and potential. There is less need than formerly to have students experience a maximum

number of grades before leaving school; however, there exists another need, to provide each student with variety, quality, success, and challenge. The data imply that some students in the upper elementary grades are denied some of these characteristics of ideal educational growth and opportunity.

5. A Cooperative Approach. In times such as the present when considerable focus is placed on evaluation and accountability, concern is expressed as to the most appropriate agent or agents to conduct evaluation studies in school systems. Arguments are put forward in favour of various groups taking responsibility for evaluation - the Ministry of Education through its central or regional offices; local school systems; families of schools; individual schools; teacher groups; research institutes; and faculties of education. Also widely discussed are proposals for cooperative evaluation, but generally neither the cooperative forms nor the functions most appropriately performed by each of the agents has been clearly delineated. The present project represents one prototype for cooperative evaluation. The study was initiated, supported, and monitored under contract by the Ministry of Education; the design, administration, and analysis were conducted by a research institute; and the local school system supported the study in a number of ways - through official support and approval, by assisting with test administration, through the help of numerous educators who served as referees and developers, and eventually through the attention the results will receive in program development.

This prototype for cooperative evaluation seems to be desirable and workable in a study of this kind where a design must be produced which takes into account a number of variables, and in which simple comparisons are therefore inadequate. When the design also creates the probability of new tests being constructed, there is obvious merit in the involvement of an institute which has the resources to evaluate the tests themselves and to design and validate comparable instruments.

Similarly, it appears reasonable to involve the Ministry in evaluation studies which may be generalizable, either in design or in potential findings useful beyond the school system in which the study is conducted. Cost factors alone dictate that each school system could not engage in an evaluation study which has as many development components as this project.

The support and cooperation of the local school system is a vital component of this model. While the individual school system probably could not afford, and in most cases does not have the concentration of research competency to conduct a study on such a massive scale, it has both the need and the setting for the cooperative study. The availability of classes, the cooperation of principals and teachers, the toleration of interruptions of programs and time-tables, and involvement with the many public relations aspects of the study represent an important contribution by the school system; in return, the greatest potential benefit of such a study accrues to the system in which it is conducted. These benefits are represented not only in the findings themselves, but also in the training and experiences acquired through the involvement of their own personnel.

The model described or referred to in this section is not considered to be the only appropriate one for cooperative evaluation. However, further replications of this particular longitudinal study, assuming that the design and tests would have broader application, could be conducted either by a school system itself, or by a regional office of the Ministry. In such replications of the study, one obvious essential would be a reduction of costs. One modification to the design that would meet this need would be the use of sampling procedures rather than procedures which tested all students on all tests.

The study demonstrated the paramount need for frank, open, and continuing communication among the cooperating agencies. The enlargement of the study beyond the minimum requirements of the contract, the construction of revised tests, and interpretations of provincial policy all created frequent needs to communicate with officials of the Ministry. Despite problems in administering the tests, open communication at all times was maintained between the researchers and the schools in the study, with each side developing higher levels of respect for and involvement with the other. In the months since the period of test administration the researchers have been asked by several of the schools for assistance with other projects or activities for which they felt they needed planning, design, or research help.

6. The Confidentiality Issue. In any study designed to respond in a comprehensive manner to public criticism about educational accountability, there is inevitably a defensive reaction on the part of those being assessed and those responsible for the program under study. Principals and teachers wonder how the data will be used; students want feedback from their tests; board officials are concerned about the ultimate findings. Those involved locally are not the only persons interested in the findings; educators and professional researchers from outside the system request, sometimes persistently, any information that may be available. The problems associated with requests for information are in some respects heightened and in others lessened if, as in the present study, maximum confidentiality must be maintained as a condition of the contract.

Tne major research advantage associated with confidentiality is freedom of feedback effects on subsequent tests; in this respect the third phase matched the first two. The original researchers withheld their findings from teachers and students until all overlapping components of the study were completed. Because of the use of revised tests in the current phase, it was important that previous tests not be taught or the answers not be "taken up" with the pupils. Similarly no results of any tests were reported back to school officials, teachers, or students. To prevent accidental disclosure, the final calculations of the tabulated data were calculated months after the tests were given and marked.

Finally, an inherent strength of the current phase of the study, both for the sake of confidentiality and from the standpoint of objectivity, was the involvement of an outside agent. No one could challenge the objectivity of the two researchers in the first two phases of this study; few boards even today could locate personnel from within their own systems who could bring both the research quality and objectivity to similar evaluations in their own schools that were displayed by Brown and Partlow. Nevertheless, from a research viewpoint the most obvious control for experimenter bias is the use of external agents.

Along with the arguments for confidentiality in a study of this kind, future investigators should be aware of two major disadvantages. Principals, teachers, and students feel cheated or short-changed when test results are withheld from them, and such lack of feedback no doubt denies teachers the opportunity to improve their own assessment, diagnosis, and re-teaching functions. The failure to provide feedback also creates unfavourable public relations for the agency involved; although this problem was anticipated by the investigator, and steps were taken, with some success, to allay criticisms and misunderstandings, future investigators planning such a massive testing program would be well advised to set up a planned series of school visits and meetings prior to the testing program, in addition to the memoranda used in this study, to ensure that principals and teachers are aware of and supportive of the nature, objectives, and activities of the project before it is administered.

7. Design For Longitudinal Studies. The most frequent and persistent criticisms of elementary schools have to do with allegations that students do not achieve as much as they did in the past, and specifically at the time the critic attended school. Debate on the substance of the argument does not appear to be very productive, because of the limited amount of objective data available for comparative purposes by those conducting the debate. What is needed is a collection of objective data about student performance in similar environments and with comparable populations over extended periods of time. In order that discussions on the question can have a clearer focus, those discussing the question need to assess whether the focus is actually on student performance, or on a disagreement about the acceptable objectives of the school.

This study represents one approach to this need, and illustrates a number of the major considerations that must be made in conducting a longitudinal study for comparative purposes. The first requirement is a comprehensive set of data for each period in question, preferably based on the same or on highly comparable or equivalent tests; if normative tests are available this aspect of the study is further enhanced. Reasonably uniform procedures for test administration, illustrated in the present project, are desirable with respect to the validity of the data. The design should also provide for prior anticipation and common treatment of problems and difficulties in communication, test administration, marking schemes, interpretation, analysis, and other factors that ensure comparability. The design should provide for means of generating new tests or new test items, and of documenting the necessity of such revisions and the means of validating them. Finally, and most important other than the use of comparable tests in the various phases of the study, is the provision for treatment of variables which themselves could account for changes in test scores. While the four variables treated in the present design - socio-economic factors, age-grade levels, vocabulary, and school objectives - may vary in importance from one phase of a study to another, or from one design to another, they represent in the present study the important characteristic of any such design, namely recognition of important factors which account for changes in student responses on the same tests administered at different times.

Any school system wishing to replicate this study for its own purposes would be able to take advantage of things learned in the present study. A number of available tests could be matched against objectives before being administered, thereby reducing the need for test construction. Reasonable sampling procedures would greatly reduce the administrative time involved. The total testing program, if carefully planned, could be administered over a more extended period of time, thereby avoiding conflicts with existing programs.

8. New Test Construction. The revised tests used in this study are themselves important contributions to evaluation, particularly in longitudinal studies. The construction of the comprehensive test

in mathematics, the mathematics sub-tests, and the vocabulary test all took into consideration modern vocabulary and school objectives. With further minor validation and possible revision they will stand as highly satisfactory tests in these subject areas. The reader will recall that all test items in the new tests were either selected from validated items in other current programs, or reflected general agreement on validity checks by panels of referees who are practising educators. It is the intention of the research team to carry out the further work needed to validate these instruments, and they will be made available to the profession in their new format. It is interesting to note that the number of students who have written these tests, and therefore for whom a data bank is available, exceeds in number the size of the population commonly employed for test validation. The one weakness of this pool, however, is that all the students are from one school system and therefore have had relatively common conditions and experiences.

9. Data Bank. The immense data bank developed in the study will become available to researchers for a wide variety of future purposes. The bank has certain advantages over many other sets of test data. For example, the tests were administered uniformly, and were not only marked uniformly but were also marked a second time and then further checked to ensure control on the marking scheme and on marking procedures. The bank provides data on student performance on precisely the same tests over a forty-year period, and also on tests reflecting vocabulary and objectives in the mid-1970's. The size of the pool, with over 10 000 test papers, tabulation sheets in the hundreds, plus the numerous tables required for analysis, presents a rich resource at three different levels, useful not only for future comparisons of a longitudinal nature, but for accurate descriptions of the period in question.

Second in importance only to generalizations about changes in student performance, at least in the view of the researcher, is the potential use of the current study as a firm baseline for comparison with future plateaus or phases in a similar or comparable study. The three phases of this study present not only objective and careful evaluations of student performance in the three time periods in question, but also probably provide the only such data banks for these purposes available in the province of Ontario. With these sources of data available, and particularly with the breadth available for the latter two phases, coupled with the treatment of major variables introduced in the second phase and more thoroughly treated in the current phase, researchers in ten or twenty years' time will be able to make an even more thorough study of student performance and the factors that affect it, and thereby continue to improve the study itself and the source of data for yet other researchers in future.

The very existence of these three studies should lend encouragement to further studies, just as Partlow was motivated by the Brown study and the present researcher by the earlier two.

Without excessive amounts of effort many school systems today could organize information about student performance for a recent period of time, or at least for the present. An initial data bank of this kind in any school system would serve a number of purposes. First of all it would enable the system to compare its achievement levels with those documented in this study. Far more important, it would provide the first baseline for a similar study within the school system. The existing study should give useful assistance to the coordinator of any such study in a school system, in such matters as test selection and revision, general design, and the treatment of major variables.

Related to the use of such a data bank is the complex issue of cybernetics. For each of the initial phases of the study, student responses to test items were appropriately recorded on answer sheets suited to hand-scoring. In the recent phase, the present investigator had to determine whether a similar marking form or a marking sheet suited to computer scoring would be more efficacious. The decision was made to remain consistent with original test conditions by having the 1975-76 subjects record responses on test answer forms that were identical to or reasonable facsimiles of the original answer forms.

Yet, in a similar future study involving the administration of over ten thousand tests, the investigator encourages future researchers to determine the quantitative difference that might be attributable to use of computer scoring sheets. Then, these differences could become adjustment ratios. The use of a computer scoring service would then be available to the investigator. With an appropriate correction factor that could be applied to account for changes due to the scoring format, comparability with existing phases could be assured. Clearly, the present investigator would have benefited in terms of time and resource utilization had such a correction factor been available in the research literature, along with instructions for modifying formats or applying the factor to various existing formats. However, given the knowledge that scoring format sheets introduce unique differences in student performance, but having no reliable means by which to quantify such differences, the research team was limited to the use of hand-scoring sheets identical or similar to those from phases one and two of the study.

Had the third phase of the study made use of computer technology, additional problems and issues would also have been raised. For example, questions requiring an "open" response would probably have needed modifications, thereby further modifying the format and altering the nature and purpose of the test. In future, the use of the computer for purposes of monitoring student achievement over time will have to be studied in the light of both fair comparison of student performance and general research design. Otherwise the resources used may result in the addition of complex extraneous variables not found in earlier phases of the study.

Bibliography

Althouse, J.G. "Structure and Aims of Canadian Education," Quance Lecture, Toronto: W.J. Gage and Co., 1949.

Anastasi, A. Psychological Testing. (3rd ed.) New York: Macmillan, 1968.

Annual Reports of Inspector of Public Schools. St. Catharines. 1933, 1934, 1935, 1936, 1939, 1941, 1942.

Annual Reports of Superintendent of Schools, St. Catharines. 1947, 1950, 1951, 1952, 1953, 1954.

Ashworth, M. A Comparative Study of Selected Background Factors Related to Achievement of 5th and 6th Grade Students. Ed.D. thesis, University of Houston, 1963.

Baratz, S.S., and Baratz, J.C. "Early Childhood Intervention: the Social Science Base of Institutional Racism." Harvard Educational Review, 40 (1970), 29-50.

Barber, T.X., and Silver, M.J. "Fact, Fiction, and the Experimenter Bias Effect." Psychological Bulletin Monograph Supplement, 70 (1968), 1-29.

Bereiter, C. "An Academically Oriented Preschool for Disadvantaged Children: Conclusions from Evaluation Studies." Hyman Bloomberg Memorial Symposium on Research in Early Childhood Education, the John Hopkins University. (February 1971).

Bereiter, C. "The Relation of Social Class and Educational Aptitude." Canadian Psychologist, 13 (October 1972), 329-40.

Bereiter, C., and Engelmann, S. Teaching Disadvantaged Children in the Preschool. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Bernstein, B. Class, Codes and Control. London: Routledge and Kegan, 1971.

Bernstein, B. "Education Cannot Compensate for Society." New Society, 26 (February 1970), 346.

Bidwell, C. "Education and Stratification," Encyclopedia of Educational Research, (4th ed.) New York: Macmillan, 1969, pp. 1247-49

Blishen, B.R.; Jones, F.E.; Naegle, K.D.; and Porter, J. (Eds.), Canadian Society: Sociological Perspectives. Toronto: Macmillan Co. of Canada Ltd., 1965.

Borus, M.E., and Nestel, G. "Response Bias in Reports of Father's Education and Socio-economic Status." Journal of the American Statistical Association, 68 (1973), 816-820.

Brodinsky, B. "Back to the Basics: The Movement and its Meaning." Phi Delta Kappan, Vol. 58 (March 1977), 522-527.

Brown, C.A. Elementary School Supervision in Ontario. Toronto: The Ryerson Press, 1949.

Boocock, S.S. "Toward a Sociology of Learning: Outside School Influences Upon Student Performance." Sociology of Education, XXXIX (Winter, 1966), 32-34.

Brookover, W. Society, Schools and Learning. Boston: Allyn & Bacon, 1969.

Campbell, D.T., and Stanley, J.C. Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally, 1966.

Caro, F.G. "Social Class and Attitudes of Youth Relevant for the Realization of Adult Goals." Social Forces, 44 (1966), 492-8.

Cary, G.L. "Class Socialization Patterns and their Relationship to Learning." School and Society, XCIV (October 29, 1966), 349-352.

Chandler, T.A. "Reading Disability and Socio-Economic Status." Journal of Reading, X (October, 1966), 5-21.

Chapin, F.S. Experimental Designs in Sociological Research. New York: Harper and Row, 1955.

Charters, W.W., Jr. "Social Class Differences in Pupil 'Success'." Handbook of Research on Teaching. (Chicago: Rand McNally, 1963), pp. 739-746.

Children and Their Primary Schools: A Report of the Central Advisory Council for Education (The Plowden Report), Central Advisory Council for Education: England. London: Her Majesty's Stationery Office, 1967.

Chopra, S.L. "Socio-Economic Background and Failure in the High School Examination." Educational and Psychological Measurement, XXVI (Summer, 1966), 495-497.

Committee on the Training of Elementary School Teachers, (The McLeod Report), Toronto: The Ontario Department of Education, 1966.

Cheal, J.E. Investment in Canadian Youth: An Analysis of Input-Output Differences Among Canadian Provincial School Systems. Toronto: The MacMillan Co. of Canada Ltd., 1963.

Coleman, J.S. Equality of Educational Opportunity. (Summary) (Washington, D.C.: U.S. Government Printing Office, 1966), pp. 14-22.

Coleman, J.S. Responsibility of Schools in the Provision of Equal Educational Opportunity. Invited Paper, Johns Hopkins University, Centre for the Study of Social Organization of Schools, ERIC Document ED 023 729, February 1968.

Coleman, J.S.; Campbell, E.Q.; Hobson, C.J.; McParland, J.; Mood, A.M.; Weinfeld, F.C.; and York, R.L. Equality of Educational Opportunity. Washington: U.S. Department of Health, Education, and Welfare, a publication of the National Center for Educational Statistics, 1966.

Demographic Influences on School Enrollment. Committee on the Costs of Education, Interim report no. 4. Toronto: Ministry of Education, 1974.

Deutsch, Martin. Minority Group and Class Status as Related to Social and Personality Factors in Scholastic Achievement. Ithaca, New York: Society for Applied Anthropology, New York State School of Industrial and Labor Relations, Cornell University, 1960. (Monograph No. 2)

Deutsch, M. "The Disadvantaged Child and the Learning Process." in Passow, A.H. (Ed.), Education in Depressed Areas, New York: Teachers College Press, Columbia University, 1963, 163-80.

Deutsch, M. "The Role of Social Class in Language Development and Cognition," American Journal of Orthopsychiatry, XXXV (January, 1965), 78-88.

Dewey, J. Democracy and Education: An Introduction to the Philosophy of Education. New York: Macmillan Co., 1955.

Drews, E.M. and Teahan, J.E. "Parental Attitudes and Academic Achievement." F. Clin. Psychol., 21 (1957), 2.

Duncan, O.D.; Featherman, D.L.; and Duncan, B. Socio-economic Background and Achievements. New York: Seminar Press, 1972.

Education. Volume 6, A Collection of Essays on Canadian Education 1964-1967. Toronto: W.J. Gage Ltd.

Education in the Primary and Junior Divisions. Toronto: Ministry of Education (Ontario), 1975.

Engelmann, S. "The Effectiveness of Direct Instruction on I.Q. Performance and Achievement in Reading and Arithmetic." In J. Hellmuth (Ed.), Disadvantaged Child, Vol. 3, New York: Brunner-Mazel, 1970, 339-361.

Finlayson, Douglas S. "Parental Aspirations and the Educational Achievement of Children." Educational Research, 14 (November 1971), 61-64.

Fleming, W.G. *Ontario's Educative Society.* Toronto: University of Toronto Press, 1972. 8 volumes.

Fleming, W.G. "Rational Strategies for Educational Change." *Anniversary Invitational Conference on Emerging Strategies and Structures for Educational Change.* Toronto: Ontario Institute for Studies in Education, June 1966.

Flesch, R. "I Say Your Child Can't Read," *Maclean's Magazine*, LXVIII (January 1, 1955), pp. 8-9.

Flesch, R. *Why Johnny Can't Read.* New York: Harper and Brothers, 1955.

Flower, G.E. "How Big is Too Big? Problems of Organization and Size in Local School Systems," *Quance Lecture*, Toronto: W.J. Gage Ltd., 1964.

Froelich, M. "Success for Disadvantaged Children." *The Reading Teacher*, Vol. 21, I (October 1967), 24.

Frost, B.P. "Some Conditions of Scholastic Achievement: Emotional and Social Adjustment." *Canadian Education and Research Digest*, V (December 1965), 278-280.

Frost, B.P. "Some Conditions of Scholastic Achievement: Part II, Environment - Home and School." *Canadian Education and Research Digest*, VI (March 1966), 5-17.

Frye, H.N. *Reports Submitted to the Joint Committee of the Toronto Board of Education and the University of Toronto.* Toronto: University of Toronto, 1962.

Giammateo, M. *Socio-economic Status and School Achievement*. Portland, Ore.: Northwest Regional Educational Laboratory, 1967.

Goldring, C.C. An Address to the Men's Club, First United Church, Port Credit (April 7, 1952), *The Toronto Globe and Mail*, April 8, 1952.

Goodlad, J.T. *et al.* *The Changing School Curriculum: A Report for the Advancement of Education.* New York: Fund for the Advancement of Education, 1966.

Goodlad, J.I., and Anderson, R.H. *The Nongraded Elementary School.* (Rev. ed.), New York: Harcourt, Brace and World Inc., 1963.

Halsey, A.H., Ed. *Education, Economy, and Society: A Reader in the Sociology of Education.* New York: Free Press of Glencoe, 1964.

Harris, R.S. *Quiet Evolution: A Study of the Educational System of Ontario.* Toronto: University of Toronto Press, 1967.

Harrison, F.I. "Opportunity As It Is Related to Home Background and School Performance." *School Review*, LXXVII (June 1969), 144-51.

Harvie, M., and Todds, P. "We Say Children Read Better," *Maclean's Magazine*, LXVIII (January 1, 1955), p. 9.

Hastings, D.E. *A Study to Determine if the Reading Comprehension of Today's Grade 6 Pupils, 1971-72, is Comparable to that of Grade 6 Pupils of 1933-38 and 1952-54 in the Elementary Schools in the City of St. Catharines.* Master's Thesis. Lewiston, N.Y.: Niagara University, 1972.

Hays, W.L. *Statistics for Psychologists.* New York: Holt, Rinehart, and Winston, 1963.

Hodgkins, B.J., and Herriott, R.E. "Social Class Structure and Education," *Encyclopedia of Education*, New York: Macmillan and Free Press, VIII (1971), 240-45.

Hope, J.A. *et al.* *The Report of the Royal Commission on Education in Ontario.* Toronto: Baptist Johnston, Printer to the King's Most Excellent Majesty (1950), XXIV - 933.

Hume, W.E. "Are the Schools Ruining Your Child?" *MacLean's Magazine*, LXV (March 1, 1952), p. 12.

Illing, W.M., and Zsigmond, Z.E. *Enrolment in Schools and Universities, 1951-52 to 1975-76.* Staff Study No. 20, Prepared for the Economic Council of Canada, Ottawa: Queens Printer, 1967.

Jantz, Richard K. "The Effects of Sex, Race, I.Q., and S.E.S. on the Reading Scores of Sixth Graders for Both Levels and Gains in Performance." *Psychology in the Schools*, II (January 1974), 90-94.

Jensen, A.R. "How Much Can We Boost I.Q. and Scholastic Achievement?" *Harvard Educational Review*, 39 (1969), 449-483.

Jensen, A.R., and Rohwer, W.D., Jr. *Experimental Analysis of Learning Abilities in Culturally Disadvantaged Children.* Final report on O.E.O. Project No. 2404, Washington, D.C.: U.S. Office of Economic Opportunity, 1970.

Jensen, A.R., and Fredriksen, J. "Social Class Differences in Free Recall of Categorized and Uncategorized Lists." In A.R. Jensen and W.E. Rohwer, Jr., *Experimental Analysis of Learning Abilities in Culturally Disadvantaged Children.* Final report on O.E.O. Project No. 2404, Washington, D.C.: U.S. Office of Economic Opportunity (1970), 103-118.

Kahl, J.A. "Educational and Occupational Aspirations of 'Common Man' Boys", Harvard Educ. Rev., XXII (1953), 186-204.

Katz, S. "The Crisis in Education," Maclean's Magazine, 66 (March 1, 1953), pp. 7-9, 54-59.

Katz, S. "The Row Over the 3 R's," Maclean's Magazine, 66 (March 15, 1953), pp. 20, 21, 47-51.

Kennett, K.F. "Intelligence and Socio-economic Status in a Canadian Sample." Alberta Journal of Educational Research, 18 (March 1972), 45-50.

Klass, A.A. "What is a Profession?" Canadian Medical Association Journals, 85 (Sept. 1961).

Laycock, S.R. "Special Education in Canada," Quance Lecture, Toronto: W.J. Gage Ltd., 1963.

Lavin, D.E. The Prediction of Academic Performance, New York: Russell Sage Foundation, 1965.

Leddy, J.F. "The Humanities in Modern Education," Quance Lecture, Toronto: W.J. Gage Ltd., 1965.

Lesser, G. "Mental Abilities of Children from Different Social Class and Cultural Groups." Society for Research in Child Development Monographs, XXX, Number 4, 1965.

Lesser, G.S.; Fifer, G.; and Clark, D.H. "Mental Abilities of Children from Different Socio-Class and Cultural Groups." Monogr. Soc. Res. Child Development, XXX, No. 7, 1965.

Living and Learning, The Report of the Provincial Committee on Aims and Objectives of Education in the Schools of Ontario (Hall-Dennis Report), Toronto: Ontario Department of Education, 1968.

Lord, F.E., and Isenberg, R.M. (Eds.) Cooperative Programs in Special Education, The Council for Exceptional Children, Washington, D.C., 1964.

Lovell, K., and Woolsey, M.E. "Reading Disability, Non-Verbal Reasoning, and Social Class." Educational Research, VI (June, 1964), 226-229.

Marjoribanks, K. "Ethnic and Evnironmental Influences on Levels and Profiles of Mental Abilities." Ed.D. thesis. University of Toronto, 1970.

Massad, C.E. "Language - Thought Processes in Children from Differing Socio-economic Levels," International Reading Association Conference Proceedings, Vol. 13, Pt. I (1969) pp. 744-48.

Meehl, P.E. "Nuisance Variables and the ex post facto Design," In M. Radner and S. Winokur (Eds.) Minnesota studies in Philosophy of Science. Minneapolis: University of Minnesota Press, 1970.

Miles, M.B. Innovation in Education. New York: Teacher's College, Columbia University, 1964.

Miller, R.I. Perspectives on Educational Change. New York: Appleton - Century - Crofts, 1967.

Mussen, P.M. et al. Child Development and Personality. New York: Harper and Row, 1956.

Napoli, J. "Environmental Factors and Reading Ability." The Reading Teacher, Vol. 21, 6 (March, 1968), 555.

National Education Association of the U.S. Project on the Instructional Program of the Public Schools, Deciding What To Teach. Washington, D.C., 1963.

Newnham, W.T., and Nease, A.S. The Professional Teacher in Ontario (Rev. ed.), Toronto: The Ryerson Press, 1967.

"Nostalgia's Child: Back to the Basics," Editorial, Phi Delta Kappan, Vol. 58, No. 7 (March, 1977), 521-22.

Nunnally, J.C. Psychometric theory. New York: McGraw-Hill, 1967.

Ontario, Department of Education.
- Intermediate Division Outlines of Courses for Experimental Use, Curriculum I:1. 1950 and 1951.
- "Physical Education" Curriculum J.5., Junior Divison: Grades 4, 5, 6. 1959
- Programme of Studies for Grades I to VI, Ontario, 1937 and 1941.
- Programme of Studies for Grades VII and VIII, Ontario 1938 and 1942.
- Programme of Studies for Grades I to 6 of the Public and Separate Schools, 1937.
- Programme of Studies for Grades I to VI of the Public and Separate Schools, 1941. Circular Elementary No. 1, 1952.
- Interim Revisions P_1J_1. Series 1966-1967.
- "Reading List" Curriculum P. and J.4: Primary and Junior Divisions, 1960.
- Reorganization of School Jurisdictions in the Province of Ontario, Jan. 1968.
- Report of the Grade 13 Study Committee, 1964.
- Reports of the Minister of Education, 1950-1967.
- Report of the Minister's Committee on the Training of Elementary School Teachers, 1966.
- Report of the Minister's Committee on the Training of Secondary School Teachers, 1962.
- "Science" Curriculum I:I (e), Intermediate Division: Grades 7, 8, 9, and 10. (To replace existing courses in Science in Curriculum I:I)

First printed for introduction in Grades 7, 8, and 9 in September, 1961 and in Grade 10 in September, 1962.
- Site: Principles of Site Development. Elementary Schools K-6, Prepared by Michael Hugh Associates Ltd., in conjunction with Ontario Department of Education.

Ontario Federation of Chapters of the Council for Exceptional Children, *Brief to the Minister's Committee on the Training of Elementary School Teachers.* Toronto: 1964.

Ontario, Legislative Assembly.
- Bill 54. An Act to amend the Public Schools Act, 1964.
- Bill 44. An Act to amend the Secondary Schools and Boards of Education Act, 1968.
- Public Enquiries Act, Revised Statutes of Ontario, 1960, c.323.
- Revised Statute of Ontario, 1960, c.330. Amendments 1964, c.95.
- The Department of Education Act.
- The Public Schools Act.
- The School Administration Act.
- The Secondary Schools and Boards of Education Act.
- The Separate Schools Act, Revised Statutes of Ontario, 1960. Toronto: Queen's Printer, 1967.
- The Teaching Profession Act, Statutes of Ontario, 1944, c.44.

Ontario, Ministry of Education. Provincial Curriculum Policy for the Primary and Junior Divisons of the Public and Separate Schools of Ontario, 1975.
- Circular P^{1}J^1, "The Formative Years."
- "Education in the Primary and Junior Divisons."

Ontario, Ministry of Government Services.
- *The Learning Society*, 1972.

Organization for Economic Cooperation and Development, *Curriculum Improvement and Educational Development*. Paris, O.E.C.D. London: Her Majesty's Stationery Office, 1967.

Partlow, Hugh R. *A Comparison of St. Catharines Public School Standards in Arithmetic and Reading 1933-38 and 1952-54.* Doctoral Thesis Abstract. Toronto: University of Toronto, 1955.

Partlow, Hugh R. *A Comparison of St. Catharines Public School Standards in Arithmetic and Reading 1933-38 and 1952-54.* Doctoral Thesis. Toronto: University of Toronto, 1955.

Partlow, Hugh R. *Arithmetic and Reading Yesterday and Today.* Toronto: The Copp Clark Co. Ltd., 1955.

Paton, J.M. Current Thinking on Teacher Education. Education, Vol. 5A. Toronto: W.J. Gage Ltd., 1966.

Phillips, C.E. The Development of Education in Canada. Toronto: W.J. Gage and Co. Ltd., 1957.

Phillips, C.E. "Public Secondary Education in Canada," Quance Lecture, Toronto: W.J. Gage and Co. Ltd., 1955.

Pitt, C.C. "Acceleration and Retardation." Canadian School Journal, 31 (August-September, 1953), 216-17.

Pitt, C.C. "Analysis of Differences In Any Given Class of Pupils." Canadian School Journal, 32 (June-July, 1954), 167-68.

Pitt, C.C. "A Philosophy of Education." School Progress, 19 (Winter, 1950), 26-27.

Porter, J. The Vertical Mosaic. Toronto: University of Toronto Press, 1965.

Quick, E.J. (Ed.) New Opportunities for the Culturally Disadvantaged. Toronto: Canadian Education Association, 1965.

Reisman, F. The Culturally Deprived Child. New York: Harper and Row, 1962.

Reller, J.L., and Morphet, E.L. (Eds.) Comparative Educational Administration. Englewood Cliffs: New Jersey, Prentice-Hall, Inc., 1962.

Report of the Ontario Legislature's Select Committee on Youth. Kingston: Hanson and Edgar Ltd., 1967.

Rist, R. "Student Social Class and Teacher Expectations." Harvard Educ. Rev., 40 (1970), 411-51.

Rosen, B.C. "The Achievement Syndrome: a Psycho-cultural Dimension of Social Stratification." Amer. Sociol. Rev., 21 (April, 1956), 203-11.

Rosenthal, R. Experimenter Bias in Behavioral Research. New York: Appleton-Century-Crofts, 1966.

Rosenthal, R., and Jacobson, L. "Teachers' Expectancies: Determinance of Pupils' I.Q. Gains." Psychological Reports, Vol. 19 (1966), No. 1.

Ryan, D.W., and Greenfield, T.B. *Clarifying the Class Size Question, Evaluation and Synthesis of Studies Related to the Effects of Class Size, Pupil-Adult, and Pupil-Teacher Ratios.* Toronto: O.I.S.E., 1976.

Scarfe, N.V. "The Aims of Education in a Free Society," *Canadian Conference on Education*, Montreal, 1962.

SEF Report E.I, "Educational Specifications and User Requirements for Elementary (K-6) Schools." *The Metropolitan Toronto School Board Study of Educational Facilities.* Toronto: Ryerson Press, 1968.

Sewell, W.H., and Hauser, R.M. *Education, Occupation, and Earnings: Achievement in the Early Career.* New York: Academic Press, 1975.

Spencer, H. *Education: Intellectual, Moral and Physical.* New York: Appleton-Century-Crofts, 1862.

Squibb, P.G. "Education and Class." *Educational Research*, 15 (June, 1973), 194-208.

Swift, D.F. "Social Class and Achievement Motivation." *Educational Research*, VIII (February, 1966), 83-95.

Tewksbury, J.L. *Nongrading in the Elementary School.* Columbus, Ohio: C.E. Merrill Books, Inc., 1967.

Universal Declaration of Human Rights. *A Guide for Teachers.* UNESCO, Paris, 1953.

Vernon, P.E. *Intelligence and Attainment Tests.* London: University of London Press, 1960.

Webb, E.J.; Campbell, D.T.; Schwartz, R.D.; and Sechrest, L. *Unobtrusive Measures: Nonreactive Research in the Social Sciences.* Chicago: Rand McNally, 1966.

Wees, W.R. "The Way Ahead," *Quance Lecture*, Toronto: W.J. Gage Ltd., 1967.

Whitehead, A.N. *The Aims of Education and Other Essays.* New York: Macmillan Co., 1959.

Wilson, Alan B. "Social Statification and Academic Achievement." *Education in Depressed Areas*. Work Conference on Curriculum and Teaching in Depressed Urban Areas, Columbia University, 1962. New York: Teacher's College Press, 1963.

Winer, B.J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill, 1962.

Wynn, J.S. "A Beginning Reading Program for the Deprived Child." *The Reading Teacher*, Vol. 21, 1, (October, 1967), 4.

Appendices

APPENDIX A: THE ARITHMETIC TESTS

A1: Arithmetic Test of Fundamental Operation for Grades IV to VIII - Form A, (The Dominion Tests)/*240*

A2: Administration Bulletin: Niagara Mathematics Test of Fundamental Operations for Grades 5 to 7/*246*

A3: Test of Fundamental Operation for Grades 5 to 7: The Niagara Mathematics Test Series, (H. & M. Revision)/*247*

A4: Group Achievement Test in Arithmetic Computation - Form A. (The Dominion Tests)/*259*

A5: Group Achievement Test in Arithmetic Reasoning - Form A, (The Dominion Tests)/*264*

A6: Group Achievement Test in Arithmetic Computation - Form B, (The Dominion Tests)/*268*

A7: Group Achievement Test in Arithmetic Reasoning - Form B, The Dominion Tests)/*272*

A8: Administration Bulletin: Niagara Mathematics Tests of Computation and Reasoning for Grade 8/*276*

A9: The Test of Arithmetic Computation for Grade 8: The Niagara Mathematics Test Series, (H.& M. Revision) /*277*

A10: Test of Arithmetic Reasoning for Grade 8: The Niagara Mathematics Test Series, (H. & M. Revision)/*282*

APPENDIX B: THE LANGUAGE ARTS TESTS

B1: The Thorndike-McCall Reading Scale for the Understanding of Sentences - Form 2/*291*

B2: Group Achievement Test in Paragraph Reading - Form A, (The Dominion Tests)/*300*

B3: Group Achievement Test in Paragraph Reading - Form B, (The Dominion Tests)/*304*

B4: Administration Bulletin: Vocabulary Test #1, 90 Common Names/*308*

B5: Vocabulary Test #1, 90 Common Names/*310*

B6: Administration Bulletin: Vocabulary Test #1, 90 Common Names (Niagara Test Series)/*318*

B7: Vocabulary Test #1, 90 Common Names (Niagara Test Series)/*320*

Appendix A1

THE DOMINION TESTS

ARITHMETIC

(FUNDAMENTAL OPERATIONS FOR GRADES IV TO VIII)

FORM A

Name .... Boy or Girl ....
(In capitals) Last First

Age .... Birthdate ....
Month Day of Month

Name of School .... Grade ....

City, Town or Municipality .... Province ....

Name of Teacher .... Date of Examination ....

INSTRUCTIONS TO PUPILS

This test contains five pages of Arithmetic examples. You are not expected to get all of the examples, but start with the first and work steadily on. Try to get as many done as you can without making mistakes. If you come to one that is too hard for you, skip it and come back to it later if you have time.

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO.

Here is a sample to show you now the examples are arranged.

1. Add 24 35 ☐	2. Subtract 12 8 ☐	3. 25 ÷ 5 ☐

You are told what to do in each example. You may figure anywhere except in the little boxes which are for your answers. Look at example number 1, above: you are told to add 24 and 35. The answer is 55, so 59 has been put in the box. Now look at number 2, in which you are told to subtract. Since 8 from 12 leaves 4, the answer 4 should go in the second box. Everyone put it in now (Pause). Now write the answer for example number 3 in the box (Pause). Since 25 ÷ 5 = 5, you should have the figure 5 in this third box. (Examiner points to box.)

In every example you do, be sure to put your answer in the box.

All answers in fractions must be given in lowest terms and improper fractions must be changed to mixed numbers in order to be counted correct.

You will have exactly 35 minutes for the test. See how many examples you can get right in that time. Remember there are five pages.

Be sure you do what you are told to do in each example.
Do not waste time on examples that are too hard for you.

Score Number Correct= ☐

1. 8+9 =

2. Add
16
48
=

3. Subtract
14
9
=

4. Multiply
9
7
=

5. 48÷6 =

6. Add
38092
4742
=

7. Subtract
100
45
=

8. Subtract
1626
829
=

9. Multiply
906
8
=

10. Add
47
53
89
=

11. Divide
9)46
=

12. Multiply
620 X 60
=

13. Add
21
846
154
1078
1397
=

14. Divide
4)3416
=

15. Multiply
381
307
=

16. Subtract
641 - 68
=

17. Divide
37)85359
=

18. Multiply
643
692
=

19. Subtract
5603
2098
=

20. Divide
93)27005
=

Go right on to Page 2

Number correct Page 1 (   )

NOTE: Be Sure to Change Improper Fractions to Mixed Numbers and Reduce all Fractions to Lowest Terms.

Weights and Measures	Fractions (A)	Fractions (B)
21. How many feet and inches in 53 inches? = ft. and in.	22. Subtract $6\frac{2}{3}$ $3\frac{1}{3}$ =	23. $\frac{1}{5}$ of 140 =
24. Subtract \$894.00 - \$31.35 =	25. $\frac{3}{8} + \frac{1}{2}$ =	26. $\frac{4}{7} \times \frac{2}{3}$ =
27. Add 3 doz. and 4 eggs, to 5 doz. and 10 eggs = doz. and eggs	28. $6 - 3\frac{5}{4}$ =	29. $\frac{1}{4} \div \frac{3}{8}$ =
30. Subtract 8 ft. 3 in. 3 ft. 10 in. = ft. in.	31. $\frac{5}{6} - \frac{2}{5}$ =	32. $3\frac{1}{3} \times 4\frac{1}{2}$ =
33. 11 times 5 pints = qt. and pt.	34. $3\frac{3}{5} + 1\frac{4}{5}$ =	35. $\frac{7}{8} \div 4$ =

Go right on to Page 3.

Number correct Page 2 (   )

Decimals (A)	Decimals (B)	Percentage
36. Copy in a Column and add 1.1 + 30. + .003 =	37. Multiply .32 .3 =	38. Change 25% to a fraction in lowest terms. =
39. Subtract 58.65 - 2.3 =	40. Multiply .12 X .12 =	41. 3% of 500 =
42. Subtract 7. .326 =	43. Place the decimal point in the product in the correct place 436 x .6 = 2616	44. $\frac{2}{3}$ is what percent? =
45. Add .23 + 1.4 =	46. Place the decimal point in the quotient in the correct place $534\overline{)\ 10.8402}$ with 203 above	47. 12 $\frac{1}{2}$ % of 800 =
48. Subtract 2.43 - .2 =	49. Divide $.04\overline{)\ 16}$ =	50. 7% of $30.00 =

Go right on to Page 4

Number correct Page 3 (   )

NOTE: Be sure to Change all Improper Fractions to Mixed Numbers and Reduce all Fractions to Lowest Terms.

Weights and measures	Fractions (A)	Fractions (B)
51. A floor is 30 ft. long and 30 ft. wide. Its area is ......sq. ft. = sq. yd.	52. $\frac{3}{4} + \frac{1}{6} + \frac{7}{12}$ =	53. $1\frac{2}{3} \div \frac{1}{4}$ =
54. Subtract 20 min. 15 sec. from 50 min. = min. sec.	55. [illegible] $\frac{1}{6} - 3\frac{4}{7}$ =	56. Change to a mixed number in lowest terms $\frac{38}{8}$ =
57. Multiply 3 yd. 2 ft. by 5. = yd. ft.	58. Add $22\frac{3}{4}$ 29 [illegible] =	59. $2 \times 1\frac{7}{8}$ =
60. Divide 3 lb. 4 oz. of tea by 2 = lb. oz.	61. Subtract $7\frac{3}{4}$ $4\frac{5}{6}$ (Answer in lowest terms) =	62. $\frac{7}{10} \div 2\frac{2}{3}$ =
63. Find the average 423 56 7 314 =	64. $31\frac{1}{8} - 5\frac{2}{9}$ =	65. $3\frac{1}{3} \div 4\frac{1}{8}$ =

Go right on to Page 5

Number correct Page 4 (   )

Percentage	Decimals and Percentage	
66. Change $37\frac{1}{2}\%$ to a fraction in lowest terms. =	NOTE: Leave these two columns until the last.	
	71. Write as decimals and add: 7 thousandths = 3 hundredths = sum =	72. $\frac{3}{4}$ of $16\frac{2}{3}\%$ of \$300 =
67. Change 5% to a decimal fraction. =	73. $1\frac{1}{4} \times 20.4$ =	74. $4.5 \times 12\frac{1}{2}\%$ of \$160 =
68. 7% of \$3.00 =	75. $6.48 \div 3\frac{3}{7}$ =	76. What part of a yard is 30 inches? (Give answer in the form of a percentage) =
69. 3 is what percent of 4? =	77. $2.75 + 4\frac{1}{5} - 3.05$ (Give answer in the form of a decimal) =	78. $\frac{1}{4}$ of 80% of .2 (Give answer in the form of a decimal) =
70. $83\frac{1}{3}\%$ of \$1200. =	79. $\frac{3.6 \times .495 \times .01}{.5 \times .004 \times 9.9}$ =	80. 3.25 min. + 45 sec. + $4\frac{1}{2}$ min. - 30 sec. = min.

THE END

If you have finished the test look over your work to see that it is correct.

Number correct Page 5 (   )

Appendix A2

Administration, Niagara Mathematics Test

DATE: MAY 31/76

TO THE TEACHER:

Re: *Administration of the Fundamental Operations Test for Grades 5 to 7.*

Please distribute onetest blank to each student and direct them to complete the information asked for on the cover page (name, grade, date, etc.). Special note should be made of the Grade 6 School section required for all Grade 7 students.

Students may use pen or pencil, and either write or print this data. When the children have completed the pupil identification section, they should be instructed to put down their pens and pencils and read the section on the front page labelled "DIRECTIONS". At this time the teacher should read the directions aloud to the class to ensure student understanding of the test format and method of response.

When the directions are completed, the students should turn immediately to "page 1", and begin answering the questions.

This test was designed to apply directly to the current mathematics curriculum in Grades 5 through 7 in Lincoln County.

The maximum testing time is 60 minutes (1 hour) but it has been found that only Grade 7 students may require more than 45 minutes to answer all 80 questions.

N.B.: Once the class has completed the test, please complete the information on the front of the test package. ALL EXTRA TEST BLANKS MUST BE RETURNED WITH THE COMPLETED TESTS. All testing should be completed by JUNE 16, 1976.

If there are any problems with test administration or test construction, please contact the O.I.S.E. office immediately -- 684-8558.

Your cooperation is sincerely appreciated by the Ministry of Education.

(H. & M. Revision)

THE NIAGARA MATHEMATICS TEST SERIES

Fundamental Operations

Grades 5 to 7

Name__________________________ Boy or Girl____________________

Age____________________ Birthday______________________________
(Month) (Date) (Year)

Grade________________ Name of School______________________________

Date____________________ Grade 6 School______________________________
(For Grade 7 students only)

DIRECTIONS: This is a test to find out how accurately you can do problems in Arithmetic.
In questions *#1 through #40* do any figuring in the blank space beside or beneath each problem.
When you have solved the problem record your answer clearly in the space for the problem.
In questions *#41 through #80* do any figuring beside the problem or on the back of the opposite page. To indicate your answer, either circle one of the choices provided or record your answer clearly in the space for the problem.

N.B.: Try to figure out each answer the best way you can, but do not spend too much time on any one question.
If you do not know anything at all about a question, leave it out and go on to the next question.

1. $7 + 8 =$

2. Add
$$\begin{array}{r} 18 \\ \underline{46} \end{array}$$

3. Subtract
$$\begin{array}{r} 16 \\ \underline{7} \end{array}$$

4. Multiply
$$\begin{array}{r} 8 \\ \underline{6} \end{array}$$

5. $45 \div 5 =$

6. Add
$$\begin{array}{r} 36081 \\ \underline{2731} \end{array}$$

7. Subtract
$$\begin{array}{r} 100 \\ \underline{55} \end{array}$$

8. Multiply
$$\begin{array}{r} 908 \\ \underline{6} \end{array}$$

9. Add
$$\begin{array}{r} 57 \\ 43 \\ \underline{69} \end{array}$$

10. Subtract
$$\begin{array}{r} 1535 \\ \underline{738} \end{array}$$

11. Multiply 260×80

12. Add
$$\begin{array}{r} 42 \\ 795 \\ 137 \\ 1096 \\ \underline{1579} \end{array}$$

(Turn to Page 2)

13. Divide $5\overline{)46}$

14. Multiply $\begin{array}{r}641\\ \underline{609}\end{array}$

15. Subtract 461 - 86

Answer = ______

16. Divide $41\overline{)94587}$

17. Multiply $\begin{array}{r}625\\ \underline{674}\end{array}$

18. Subtract $\begin{array}{r}7801\\ \underline{4096}\end{array}$

19. Divide $82\overline{)23834}$

20. $\frac{1}{7} \times 182 =$

21. \$672.00 - \$41.65 =

22. $\frac{5}{6} + \frac{1}{2} =$

(Turn to Page 3.)

______

23. $\frac{5}{6} + 3\frac{1}{2} =$

24. $\frac{5}{9} - \frac{2}{9} =$

25. $\frac{3}{4} \times \frac{2}{7} =$

26. $8 - 4\frac{2}{3} =$

27. $5\frac{5}{9} + 3\frac{1}{3} =$

28. $0.6 + 0.7 =$

29. $5.61 - 0.09 =$

30. Copy in a Column and Add:

$3.3 + 50 + 0.005$

Answer = ______________

31. Multiply $\begin{array}{r} 0.74 \\ \underline{0.7} \end{array}$

32. Change 75% to a fraction in lowest terms.

Answer = ______________

(Turn to Page 4.)

33. $\frac{2}{3} + \frac{5}{6} + \frac{1}{12} =$

34. Add 0.47 + 2.8

Answer = ____________

35. Add 6 doz. and 9 eggs
to 2 doz. and 5 eggs.

Answer = _____doz. and _____ eggs

36. Subtract 36.85 - 4.6

37.

Change $\frac{49}{9}$ to a mixed number in lowest terms.

Answer = ____________

38. 2.7 X 7 =

39. Subtract 4.83 - 0.4 =

40. How many feet and inches equal 53 inches?

Answer = _____ft. and _____ in.

(Turn to Page 5.)

Complete the Chart:

	Thousands	Hundreds	Tens	Ones		Numeral
41.	▨	5	24	12	=	______
42.	6	10	0	2	=	______
43.	______	17	9	3	=	3793

44. Which expression is equivalent to 0 X 5 ?

(Circle the answer)

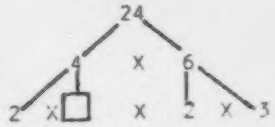
(a) 78 X 0

(b) 5 X 1

(c) 5 X 10

(d) 78 ÷ 78

45.



What factor should go in the box?

(Circle the answer)

(a) 1 (b) 2

(c) 3 (d) 4

46. Which of the following is a factor of 15 ?

(Circle the answer)

(a) 5 (b) 7

(b) 10 (d) 30

47. Which of the following is a prime number?

(Circle the answer)

(a) 15 (b) 12

(c) 11 (d) 10

48. Find the greatest common factor of 30 and 18.

(Circle the answer)

(a) 3 (b) 6

(c) 9 (d) 15

(Turn to Page 6.)

49. Use one of the following signs to make the statement true:

$$<,\ >,\ \text{or}\ =$$

$(3 + \triangle) + 5$ ______ $3 + (\triangle + 5)$

50. Which set contains only factors of both 16 and 24 ?
(Circle the answer)

(a) $\{0, 8, 16, 24\}$
(b) $\{2, 4, 6, 8\}$
(c) $\{2, 4, 8\}$
(d) $\{2, 3, 4\}$

---

51. $59 \overline{)\ 41700}$

The quotient is between
(Circle the answer)

(a) 500 and 599
(b) 600 and 699
(c) 700 and 799
(d) 800 and 899

52. The chart below shows the rainfall in Bay City.

October	$3\frac{2}{5}$ in.
November	$4\frac{3}{10}$ in.
December	$4\frac{1}{5}$ in.

What was the total rainfall for the three months?

Answer = ____________

---

53. A packing box can hold 48 books. About how many boxes will be needed to pack the 3015 books in a library?
(Circle the answer)

(a) 150,000 (b) 15,000

(c) 600 (d) 60

54. Here is a list of the number of stamps each of three boys has:

Jim 15
Frank 34
Bob 80

Bob has more stamps than Jim and Frank together. How many more?

Answer = ____________

(Turn to Page 7.)

55. A soft drink carton contains 12 cans of orange drink. The empty carton alone weighs 500 grams and each can weighs 300 grams.
What is the total weight of the full carton?
(Circle the answer)
(a) 812 grams
(b) 4100 grams
(c) 6300 grams
(d) 9600 grams

56. On a jet flight which took 6 hours the average speed was 550 m.p.h. The return flight took 5 hours.
What was the average speed on the return flight?
(Circle the answer)
(a) 550 m.p.h.
(b) 600 m.p.h.
(c) 660 m.p.h.
(d) 700 m.p.h.

57. Subtract

$$\begin{array}{r} 8\frac{3}{5} \\ 5\frac{1}{5} \\ \hline \end{array}$$

58. Divide

$$8\overline{)\,6832}$$

59. $\frac{1}{2} - 2 =$

60. $10 \times \frac{2}{3} =$

61. $3\frac{1}{2} \times \frac{5}{6} =$

62. Place the decimal point in the product in the correct place.

$872 \times 0.8 = \boxed{6\ 9\ 7\ 6}$

(Turn to Page 8.)

63. Mary had 4 yards of material. She used $1\frac{1}{2}$ yards to make a skirt and $2\frac{1}{4}$ yards for a matching jacket. How much material was left over?

(Circle the answer)

(a) $\frac{1}{4}$ yd. (b) $1\frac{1}{4}$ yd. (c) $1\frac{3}{4}$ yd.

(d) $2\frac{1}{2}$ yd. (e) $3\frac{3}{4}$ yd.

64. What is the missing numerator?

$$\frac{2}{3} = \frac{?}{18}$$

(Circle the answer)

(a) 6

(b) 12

(c) 9

(d) 36

65. Which of these statements is true?

(Circle the answer)

(a) $\frac{17}{4}$ is less than $4\frac{1}{4}$

(b) $\frac{17}{4}$ is equal to $4\frac{1}{4}$

(c) $\frac{17}{4}$ is greater than $4\frac{1}{4}$

66. Subtract 64¢ from 5 dollars.

Answer = ______________

67. Rounded to the nearest tenth, 35.749 becomes:

(Circle the answer)

(a) 35.7 (b) 35.75

(b) 35.8 (d) 36

68. Which of these statements is true?

(Circle the answer)

(a) $\frac{1}{5} = 0.2$

(b) $\frac{1}{5}$ = is less than 0.2

(c) $\frac{1}{5}$ = is more than 0.2

(Turn to Page 9.)

69. What is 8.7% of 125 ?

(Circle the answer)

(a) 16.25 (b) 10.875

(c) 108.75 (d) 162.5

70. What is 120% of 45 ?

(Circle the answer)

(a) 9 (b) 56

(c) 54 (d) 65

71. 0.05 x 0.5 =

72. 0.48 - 0.3 =

73. Find the lowest common multiple of 12 and 15.

(Circle the answer)

(a) 3 (b) 12

(c) 60 (d) 180

74. How would you write 27 hundredths as a decimal?

Answer = ____________

75. A car goes 22.3 miles on a gallon of gas. About how many gallons of gas will be needed for a trip of 200 miles?

(Circle the answer)

(a) 10 (b) 45

(c) 180 (d) 445

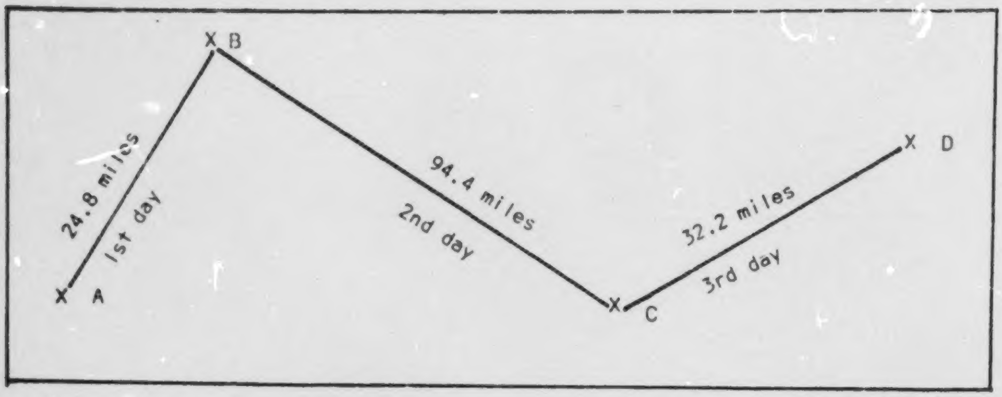
76. About 0.8 of the weight of an orange is juice. The juice is removed from 49 kilograms of oranges.
Which is the best estimate of the weight of this juice?

(Circle the answer)

(a) 4 kg (b) 32 kg

(c) 40 kg (d) 50 kg

(Turn to Page 10.)


X B
24.8 miles
1st day
X A
94.4 miles
2nd day
X C
32.2 miles
3rd day
X D

Note: The above chart shows distances Mr. Smith travelled on a three-day trip. Use it to answer questions #77 and #78.

77. What was the total distance Mr. Smith travelled in three days?

(Circle the answer)

(a) 151.4 miles

(b) 150.6 miles

(c) 140.4 miles

(d) 126.6 miles

(e) 119.2 miles

78. How many more miles did Mr. Smith travel on the second day than on the first day?

(Circle the answer)

(a) 69.6 miles

(b) 62.2 miles

(c) 57.0 miles

(d) 119.2 miles

(e) 126.6 miles

(Turn to page 11.)

79. Butter costs 69 cents a pound.

How many pounds can you buy for \$4.15?

(Circle the answer)

(a) 0.06

(b) 0.5

(c) 5

(d) 6

---

80. The 32 pupils in Mr. Brown's class are going on a picnic. Mr. Brown is going to take groups of 5 pupils at a time in his car.

The *number of trips* he will make in order to take all of the class is

(Circle the answer)

(a) 6

(b) 6 R 2

(c) $6\frac{2}{5}$

(d) 7

N.B.: *If you finish before time is up, go back and check your work!!!*

Appendix A4

Form A

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PART II

TEST 4 -- ARITHMETIC COMPUTATION

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name .... Boy or Girl ....
IN CAPITALS LAST FIRST

Age last Birthday .... Birthday ....
MONTH DATE YEAR

Grade .... Name of School ....

Province .... City, Town or Municipality ....

Date .... Name of Teacher ....

SCORES

Arithmetic Computation ...... _____

Arithmetic Reasoning .... _____

TOTAL .... _____

6. SUBTRACT $\frac{4}{5} - \frac{2}{3} =$

ANSWER ....

---

7. DIVIDE $\frac{3}{4} \div \frac{4}{5} =$

ANSWER ....

---

8. MULTIPLY $3\frac{3}{7} \times 6\frac{1}{8} =$

ANSWER ....

---

9. ADD 28.56 + 4.073 + 287.4 =

ANSWER ....

---

10. SUBTRACT 84.37 − 6.193

ANSWER ....

---

11. DIVIDE $38\overline{)\ \$27.36}$

ANSWER ....

---

12. 45% expressed as a common fraction
in its lowest terms =

ANSWER ....

---

13. $7\frac{1}{2}\%$ of 800 =

ANSWER ....

---

14. SUBTRACT $7\frac{1}{2} - 3\frac{2}{9} =$

ANSWER ....

---

15. DIVIDE $1\frac{7}{9} \div 6\frac{2}{3} =$

ANSWER ....

Continued on page 3

16. 8% expressed as a decimal =

ANSWER ....

---

17. ADD $5\frac{2}{3} + 2\frac{5}{6} =$

ANSWER ....

---

18. MULTIPLY

$$\begin{array}{r} .0\,0\,3\,4\,2 \\ \underline{3\,0.3\,1} \end{array}$$

ANSWER ....

---

19. SIMPLIFY $\frac{12}{35}$ of $\frac{14}{27} =$

ANSWER ....

---

20. MULTIPLY

$$\begin{array}{r} \$2.85 \\ \underline{5\frac{1}{3}} \end{array}$$

ANSWER ....

---

21. ADD $\frac{5}{8} + \frac{1}{6} + \frac{7}{12} =$

ANSWER ....

---

22. $\frac{8}{25}$, written as a decimal =

ANSWER ....

---

23. DIVIDE $.06\overline{)\,.0\,0\,4\,9\,2\,8}$

ANSWER ....

Continued on page 4

TEST 4 -- ARITHMETIC COMPUTATION

DIRECTIONS: This is a test to find out how rapidly and accurately you can do problems in Arithmetic. Do any necessary figuring in the blank space beside each problem, and as soon as you have finished a question, copy the answer in the place provided to the right of that question.

As there is a time limit, you will need to work as rapidly as possible. Do not waste too much time over any one question. Skip any questions that seem too difficult and return to them later if you have any time left.

---

1. ADD

$$\begin{array}{r} 77732 \\ 6958 \\ 56342 \\ 986 \\ 42504 \\ 7849 \\ \hline \end{array}$$

ANSWER ....

2. SUBTRACT

$$\begin{array}{r} 7540503 \\ 4608297 \\ \hline \end{array}$$

ANSWER ....

3. MULTIPLY

$$\begin{array}{r} 3702 \\ 7003 \\ \hline \end{array}$$

ANSWER ....

4. DIVIDE

$$273\overline{)165711}$$

ANSWER ....

5. Find the average of

$$\begin{array}{r} 8 \\ 453 \\ 76 \\ \hline \end{array}$$

ANSWER ....

Continued on page 2

24. 15 is what per cent of 25?

ANSWER ....

---

25. DIVIDE 3) 14 gal. 1 qt.

ANSWER ....

End of Test in Arithmetic Computation. Look over your work.

Pupil's score (number right) ______

Appendix A5

Group Achievement Test, Form A (The Dominion Tests)

TEST 5 -- ARITHMETIC REASONING

DIRECTIONS: This is a test to find out how well you can do problems in Arithmetic. Do any necessary figuring in the blank space beside each problem, and as soon as you have finished a question, copy the answer in the place provided to the right of that question.

As there is a time limit you will need to work as rapidly as possible. Do not waste too much time over any one question. Skip any questions that seem too difficult and return to them later if you have any time left.

---

1. What is the interest on $500 for 3 years at 8% simple interest? ANSWER ....

---

2. A certain map is drawn so that one-eighth of an inch represents a mile. How many miles apart are two towns which are 3 inches apart on the map? ANSWER ....

---

3. A man bought a house for $8,000 and sold it to gain 12%. What was his gain in dollars? ANSWER ....

---

4. A grocer bought a crate of canned salmon for $5.40 and retailed it for $7.20. How many cans did the crate hold, if his profit was 5c a can? ANSWER ....

---

5. My store is valued at $9,000. It is assessed at 2/3 of its value, and the tax rate is 30 mills. What is my annual tax bill? ANSWER ....

Continued on page 2

6. A farmer has 80 head of stock. One-tenth of them are horses, one-fourth are cattle, two-fifths are hogs and the rest are sheep. How many sheep has he?

ANSWER ....

7. An article is listed at $10.00. What do I pay for it if I buy it at discounts of 20% and 10%?

ANSWER ....

8. A man bought a house for $4,000. He spent $300 in repairs, rented it at $25 a month for 10 months and then sold it for $4,200. Did he gain or lose, and how much?

ANSWER ....

9. How many acres are there in a farm 320 rods long and 200 rods wide?

ANSWER ....

10. A man walks at the rate of 4 miles per hour. How many miles will he walk from 2:30 p.m. to 5:15 p.m.?

ANSWER ....

11. A dealer buys cars at $800 and sells them at $1,000. His gain is what per cent of the selling price?

ANSWER ....

12. Find the cost of digging a trench 120 yards long, 3 feet wide and 6 feet deep at 40c a cubic yard.

ANSWER ....

13. A druggist buys a certain tonic at $6.00 a dozen bottles and sells it at 80c a bottle. What is his gain per cent, calculated on the cost price?

ANSWER ....

Continued on page 3

14. A rectangle has an area of 24 square inches. Its width is 3 inches. What is its perimeter? ANSWER ....

15. I paid $4.00 for an article on which the marked price was $5.00. What rate of discount did I receive? ANSWER ....

16. Two boys rented a tennis-court for a month for $15, agreeing to divide the cost in proportion to the number of times each played. Jim played 12 times and Tom 8. How much of the $15 should Jim pay? ANSWER ....

17. A dealer sells a car for $1,200, making a gain of 20% on the cost price. What was the cost price? ANSWER ....

18. I make a solution of alcohol by mixing 16 pints of water with 9 pints of alcohol. What per cent of the solution is alcohol? ANSWER ....

19. A playground is 12 rods long and 8 rods wide. What will it cost to fence it at 60c a yard? ANSWER ....

20. An outdoor swimming tank 20 yards long and 15 yards wide is surrounded by a cement walk 2 yards wide. What is the area of the walk in square yards? ANSWER ....

End of Test in Arithmetic Reasoning. Look over your work.

Pupil's score (number right) ______

Form A

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PART II
TEST 5 -- ARITHMETIC REASONING

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name .... Boy or Girl .....
IN CAPITALS LAST FIRST

Age last Birthday .... Birthday ....
MONTH DATE YEAR

Grade .... Name of School ....

Province .... City, Town or Municipality ....

Date .... Name of Teacher ....

	Scores
Arithmetic Computation	______
Arithmetic Reasoning	______
TOTAL	______

FORM B

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PART II

ARITHMETIC COMPUTATION

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name ______________________ Boy or Girl ______________

Age ______________ Birthday ______________________
Month Date Year

Grade ____________ Name of School ______________________

Date ____________ Name of Teacher ______________________

	SCORES
ARITHMETIC COMPUTATION ...	
ARITHMETIC REASONING	
TOTAL	

TEST 4 -- ARITHMETIC COMPUTATION

DIRECTIONS: This is a test to find out how rapidly and accurately you can do problems in Arithmetic. Do any necessary figuring in the blank space beside each problem, and as soon as you have finished a question, copy the answer in the place provided to the right of that question.

As there is a time limit, you will need to work as rapidly as possible. Do not waste too much time over any one question. Skip any questions that seem too difficult and return to them later if you have any time left.

---

1. ADD

$$\begin{array}{r} 72354 \\ 39667 \\ 3869 \\ 506 \\ 84659 \\ \underline{47382} \end{array}$$

ANSWER ....

2. SUBTRACT

$$\begin{array}{r} 9660806 \\ \underline{5709149} \end{array}$$

ANSWER ....

3. MULTIPLY

$$\begin{array}{r} 6017 \\ \underline{9004} \end{array}$$

ANSWER ....

4. DIVIDE

$$274\overline{)192896}$$

ANSWER ....

5. Find the average of

$$\begin{array}{r} 6 \\ 587 \\ \underline{64} \end{array}$$

ANSWER ....

6. SUBTRACT

$$\frac{6}{7} - \frac{2}{3} =$$

ANSWER ....

Continue on page 3

7. DIVIDE $\frac{2}{3} \div \frac{5}{7} =$

ANSWER ....

---

8. MULTIPLY $7\frac{4}{5} \times 6\frac{2}{3} =$

ANSWER ....

---

9. ADD 34.65 + 9.087 + 160.6 =

ANSWER ....

---

10. SUBTRACT 47.52 - 2.351 =

ANSWER ....

---

11. DIVIDE

$52\overline{)\ \$33.28}$

ANSWER ....

---

12. 32%, expressed as a common fraction in its lowest terms =

ANSWER ....

---

13. $3\frac{1}{2}$% of 1200 =

ANSWER ....

---

14. SUBTRACT $8\frac{5}{6} - 5\frac{1}{4} =$

ANSWER ....

---

15. DIVIDE $1\frac{7}{25} \div 4\frac{4}{5} =$

ANSWER ....

---

16. 3%, expressed as a decimal =

ANSWER ....

Continue on page 4

17. ADD $8\frac{1}{2} + 2\frac{5}{6} =$

ANSWER ....

---

18. MULTIPLY

$$\begin{array}{r} .00472 \\ \underline{40.41} \end{array}$$

ANSWER ....

---

19. SIMPLIFY $\frac{15}{28}$ of $\frac{16}{21} =$

ANSWER ....

---

20. MULTIPLY

$$\begin{array}{r} \$3.56 \\ \underline{4\frac{1}{2}} \end{array}$$

ANSWER ....

---

21. ADD $\frac{2}{9} + \frac{5}{6} + \frac{1}{3} =$

ANSWER ....

---

22. $\frac{7}{20}$, written as a decimal =

ANSWER ....

---

23. DIVIDE $.07\overline{)\ .0005391}$

ANSWER ....

---

24. 28 is what per cent of 35?

ANSWER ....

---

25. DIVIDE $2\overline{)\ 9\text{ yds. }2\text{ ft. }8\text{ ins.}}$

ANSWER ....

End of Test in Arithmetic Computation. Look over your work.

Pupil's score (nuber right) ____________

FORM B

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PART II

ARITHMETIC REASONING

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name ______________________ Boy or Girl __________

Age __________ Birthday ______________________

Month Date Year

Grade __________ Name of School ______________________

Date __________ Name of Teacher ______________________

SCORES

TEST 5—ARITHMETIC REASONING

DIRECTIONS: This is a test to find out how well you can do problems in Arithmetic. Do any necessary figuring in the blank space beside each problem, and as soon as you have finished a question, copy the answer in the place provided to the right of that question.

As there is a time limit, you will need to work as rapidly as possible. Do not waste too much time over any one question. Skip any questions that seem too difficult and return to them later if you have any time left.

---

1. What is the interest on $600 for two years at 5% simple interest? ANSWER..................

2. A map is drawn so that 1/16 inch represents a mile. How many miles apart are two towns which are 2 inches apart on the map? ANSWER.................

3. Mr. White bought a farm for $7,000, and sold it at a loss of 15%. What was his loss in dollars? ANSWER..................

4. My house is valued at $8,000. It is assessed at ¾ of its value and the tax rate is 40 mills. What is my annual tax bill? ANSWER..................

5. A farmer took 60 sacks of potatoes to market. He sold one-third of them to Mr. Jones, one-tenth of them to Mr. Kelly, two-fifths of them to Mr. Brown and the rest to Mr. Smith. How many sacks did Mr. Smith buy? ANSWER..................

Continue on page 2

6. If a train runs at the average rate of 30 miles per hour, how many miles will it travel from 2 p.m. to 4.20 p.m.? ANSWER..................

---

7. A typewriter is listed at $75. What does it cost me if I am allowed a trade discount of 20% and a further discount of 5% for cash payment? ANSWER..................

---

8. I buy a house for $7,500 and sell it for $10,000. My gain is what per cent of the selling price? ANSWER..................

---

9. A merchant bought a shipment of peaches for $105, and retailed them for $119. If his profit was 7c a basket, how many baskets were there in the shipment? ANSWER..................

---

10. For a farm 120 rods long and 80 rods wide a man paid $75 an acre. How much did the farm cost him? ANSWER..................

---

11. A grocer buys vinegar at $1.80 a dozen bottles and sells it at 15c a bottle. What is his gain per cent calculated on the cost price? ANSWER..................

---

12. A rectangle has an area of 14 square inches. Its width is 2 inches. What is its perimeter? ANSWER..................

---

13. A customer paid $3.60 for an article on which the marked price was $4.00. What rate of discount did he receive? ANSWER..................

Continue on page 3

14. Find the cost of filling a tank 15 yards long, 3 feet wide and 6 feet deep at 25c a cubic yard.

ANSWER.................

---

15. I bought a house for $7,000. I paid $250 taxes, and after receiving a year's rent at $40 a month, sold it for $6,900. Did I gain or lose, and how much?

ANSWER.................

---

16. A fruit-grower pays $3.20 to two boys for picking apples. If Bill picks 15 baskets and Fred picks 25 baskets, how much of the money should Bill receive?

ANSWER.................

---

17. A blend of tea is made by mixing 10 lbs. of green with 15 lbs. of black tea. What per cent of the mixture is green tea?

ANSWER.................

---

18. A real estate dealer sold a house at a profit of 25% on the cost. If the selling price was $10,000, what was the cost price?

ANSWER.................

---

19. A playground is 18 rods long and 12 rods wide. What will it cost to fence it at 50c a yard?

ANSWER.................

---

20. An outdoor swimming tank 30 yards long and 20 yards wide is surrounded by a cement walk 2 yards wide. What is the area of the walk in square yards?

ANSWER.................

End of Test in Arithmetic Reasoning. Look over your work.

Pupil's score (number right)————

Appendix A8

DATE: MAY 31, 1976

TO THE TEACHER:

Re: *Administration of the Grade 8 Computation and Reasoning Tests.*
(Niagara Mathematics Test Series)

Please distribute one test blank to each student and direct him/her to complete the information asked for on the cover page (name, grade, etc.). Special note should be made of the Grade 6 School section required for all Grade 8 students.

Students may use pen or pencil, and either write or print this data. When the students have completed the pupil identification section, they should be instructed to put down their pens and pencils, and read the section on the front page labelled "DIRECTIONS".

At this time the teacher should read the directions aloud to the class to ensure student understanding of the test format and method of response.

When the directions are completed, the students should turn immediately to "page 1", and begin answering the questions.

This test was designed to apply directly to the current mathematics curriculum in Grade 8 in Lincoln County.

The maximum testing time for the Computation Test and the Reasoning Test is 30 minutes (½ hour) each. Most Grade 8 students should complete the Computation Test in 25 minutes and the Reasoning Test in 20 minutes. It is recommended that the Computation Test be given first in order to familiarize the students with test format and method of response.

N.B.: Once the class has completed the test, please complete the information on the front of the test package and note any extraordinary circumstances encountered during the test.

All extra test blanks must be returned with the completed tests. All testing should be completed by JUNE 18, 1976.

If there are any problems with test administration or test construction, please contact the O.I.S.E. office immediately -- 684-8558.

Your cooperation is sincerely appreciated by the Ministry of Education.

Appendix A9

(H. & M. Revision)

THE NIAGARA MATHEMATICS TEST SERIES

Computation

Grade 8

Name ______________________________ Boy or Girl __________

Age ____________________ Birthday ____________________

Grade ______________ Name of School ____________________

Date ______________ Grade 6 School ____________________

(For Grade 8 Students Only)

DIRECTIONS: This is a test to find out how accurately you can do problems in arithmetic.

In questions *#1 through #24* do any figuring in the blank space beside or beneath each problem.

When you have solved the problem, record your answer clearly in the space for the problem.

In question #25 do any figuring in the space provided with the problem. To indicate you answer circle one of the choices provided.

N.B. Try to figure out each answer the best way you can, but do not spend too much time on any one question.

If you do not know anything at all about a question, leave it out and go on to the next question.

1.

$$\begin{array}{r} 563 \\ 18 \\ 777 \\ 9 \\ 405 \\ 69 \\ \hline \end{array}$$

2. Subtract

$$\begin{array}{r} 75402 \\ 46085 \\ \hline \end{array}$$

3. Multiply

$$\begin{array}{r} 7103 \\ 3702 \\ \hline \end{array}$$

4. Divide

$$27\overline{)1647}$$

5. Find the average of:

5 6 and 4 7 3 and 8

6. Subtract

$$\frac{4}{5} - \frac{2}{3} =$$

7. Divide

$$\frac{4}{5} \div \frac{3}{4} =$$

8. Multiply

$$6\frac{1}{7} \times 3\frac{3}{8} =$$

(Turn to Page 2)

9. Add

$46.85 + 6.091 + 465.2 =$

10. Subtract

$62.59 - 8.375 =$

11. Express 35% as a common fraction in its lowest terms.

12. What is $5\frac{1}{2}$% of 600 ?

13. Subtract

$9\frac{1}{2} - 5\frac{3}{9} =$

14. Divide

$3\frac{5}{9} \div 9\frac{1}{3} =$

15. Express 6% as a decimal.

16. Multiply

$$\begin{array}{r} 60.62 \\ .0684 \\ \hline \end{array}$$

(Turn to Page 3)

17. Add

$\frac{5}{12} + \frac{1}{6} + \frac{3}{8} =$

18. Divide

$18 \div 6 =$

19.

$3 \times 8 \div 2 + 2 \times 3 =$

20. If $X = \frac{3}{4}$,

then $X^2 - \frac{1}{4}X$ equals __________

21.

Simplify: $\sqrt{\frac{16}{25}}$

22.

Write $\frac{4}{25}$ as a decimal.

23. Add

$7\frac{1}{3} + 4\frac{5}{6} =$

24. Divide

$57 \overline{)\$41.04}$

(Turn to Page 4)

____

25. The expression $9(10^6) + 3(10^5) + 7(10^3) + 8(10) + 6$ represents the same number as:

(Circle the Answer)

(a) 9,370,086

(b) 9,037,086

(c) 9,300,786

(d) 9,307,086

N.B.: *If you finish before time is up, go back and check your work!!!*

Appendix A10

(H. & M. Revision)

THE NIAGARA MATHEMATICS TEST SERIES

Reasoning

Grade 8

Name ________________________ Boy or Girl ______________

Age ________________ Birthday ________________________

Grade __________ Name of School ______________________

Date ________________

Grade 6 School ________________

DIRECTIONS: This is a test to find out how well you understand problems in mathematics. You should have plenty of time to at least attempt all 20 questions in this test.

In questions #1 *through* #12 read each question very carefully and do any figuring in the blank space beneath each problem.

When you have solved the problem record your answer clearly in the space provided.

In questions #13 *through* #20 read each question very carefully and do any figuring in the blank space beneath each problem.

When you have solved the problem record your answer clearly by circling one of the choices provided.

N.B. Try to figure out each answer the best way you can, but do not spend too much time on any one question. If you do not know anything at all about a question, leave it out and go on to the next question.

1. What is the interest on $500.00 for 3 years at 8% simple interest?

ANSWER____________________

---

2. A certain map is drawn so that one-tenth of an inch represents a mile. How many miles apart are two towns which are 3 inches apart on the map?

ANSWER____________________

---

3. A man bought a house for $80,000 and sold it to gain 12%. What was his gain in dollars?

ANSWER____________________

---

4. The student's council of a school received delivery of a shipment of T-shirts costing a total of $54.00. If the T-shirts were then sold for a total of $72.00, how many T-shirts were in the shipment if the student's council made a profit of 50¢ a shirt?

ANSWER____________________

(Turn to Page 2.)

____

---

6. A record store marked a $10.00 album with a tag showing a price which is 20% below the list price. During a one-day special a further discount of 10% is given on the tagged price. What price would someone pay if he/she bought the album during the special sale?

ANSWER________________

---

7. A man bought a house for $40,000. He spent $3,000 in repairs. After completing the repairs he rented the house to a family for $250 a month for a 10 month period. If he sold the house for $42,000 to a real estate company at the end of the 10 month period, how much did he gain or lose?

ANSWER________________

(Turn to Page 3.)

8. During a charity walk-a-thon a student walks at the rate of 4 miles per hour. How many miles will he walk from 1:30 p.m. to 6:15 p.m.?

ANSWER________________

---

9. A dealer buys motorcycles at $800 and sells them at $1,000. His gain is what percent of the selling price?

ANSWER________________

---

10. Find the cost of digging a swimming pool which is 40 feet long, 30 feet wide, and 6 feet deep at 50¢ a cubic foot.

ANSWER________________

---

11. A sheet of paper has an area of 88 square inches. Its width is 8 inches. What is its perimeter?

ANSWER________________

---

12. I paid $4.00 for sweatshirt on which the marked price was $5.00. What rate of discount did I receive?

ANSWER________________

(*Turn to Page 4.*)

13. The reciprocal of 78 is
(Circle the Answer)

(a) 87

(b) $\frac{8}{7}$

(c) $\frac{1}{78}$

(d) $\frac{7}{8}$

---

14. The set of odd whole numbers divisible by 2 is
(Circle the Answer)

(a) $\{0, 2, 4, 6, \ldots\}$

(b) $\{0, 2, 6, 10, \ldots\}$

(c) $\{0\}$

(d) $\{\ \}$

---

15. If n and $7.777 \times n$ are whole numbers, then the smallest value of n is
(Circle the Answer)

(a) $\frac{1000}{7777}$

(b) $\frac{1000}{777}$

(c) 10^4

(d) 10^3

(Turn to Page 5.)

16. If Johnny Bench's batting average was .250, he averaged *(Circle the Answer)*

(a) 2 hits out of every 4 times at bat

(b) 2 hits out of every 8 times at bat

(c) 4 hits out of every 12 times at bat

(d) 4 hits out of every 20 times at bat

(Turn to Page 6.)

17. In a track and field meet, 20 students competed in the races, 16 in the high jump, and 10 in the pole vault.
Of these, 3 competed in the races and the high jump,
4 competed in the high jump and the pole vault,
6 competed in the races and the pole vault.
Of the above 13 people, 2 competed in all three events.
The number that competed only in the races is

(Circle the Answer)

(a) 7

(b) 9

(c) 13

(d) 15

---

18. A tourist decided to go to Mexico for a holiday and he exchanged 100 Canadian dollars for pesos. One peso was worth 0.090 U.S. dollars and 1 Canadian dollar was worth 0.93 U.S. dollars. If there was no service charge, how many pesos did he receive, to the nearest peso?

(Circle the Answer)

(a) 837

(b) 968

(c) 1033

(d) None of the above.

(Turn to Page 7.)

19. Ten groups of 20 students were chosen at random from a large school and asked to indicate which brand of toothpaste they used. The following data were recorded.

GROUP	1	2	3	4	5	6	7	8	9	10
Number using Brand 'A'	8	7	5	16	14	10	9	8	4	9
Number using Brand 'B'	10	12	8	1	4	10	9	4	2	0
Number using Other Brands	2	1	7	3	2	0	2	8	14	11

Assuming that the groups chosen were representative of the school, what percent of all the students could be expected to use Brand 'A'?

(Circle the Answer)

(a) 25

(b) 30

(c) 45

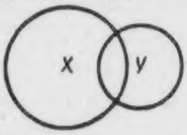
(d) 90

(Turn to Page 8.)

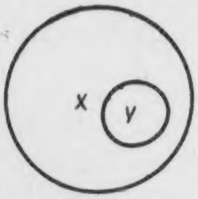
20. Which one of the following Venn diagrams illustrates that X is a subset of Y?

(*Circle the Answer*)

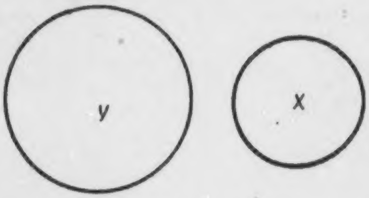
(a)



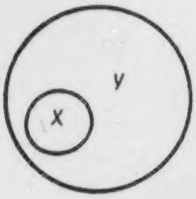
(b)



(c)



(d)



N.B. *If you finish before time is up, go back and check your work!!!*

Appendix B1

THORNDIKE-McCALL READING SCALE

For the Understanding of Sentences

FORM 2

STUDENT'S NAME ______________________________ DATE ____________

SCHOOL ______________________________ GRADE ____________

AGE ______________ BIRTH DATE ______________________

This is to be a reading contest. You will read paragraphs like this one, and answer questions like those you see below. Answer every question you can. If you come to a question you can't answer, skip it and go on. Go back to it later. If you finish before you are told to stop, go back and make sure you have made no mistakes. When possible the answers to the questions must be found in the paragraph. You may read the paragraph as many times as you need to. You will have enough time but don't waste it. Play fair. Don't look at anyone else's paper. You will be told your score later.

1. Does it say you are to do your best? ______________________
2. Does it say you may read the paragraph as often as you need to? __________
3. Are you to look at anyone else's paper? ______________________
4. When possible, where must you find the answers? ______________________

____

As soon as you finish one page do the next. You are to start with the first page. Open paper! Begin!

Read this and then write the answers. Read it again if you need to.

Nell's mother went to the store on Water Street to buy ten pounds of sugar, a dozen eggs and a bag of salt. She paid a dollar in all. Nell and Joe went with her. On the way home on Pine Street, they saw a fire-engine with three horses.

1. Was the salt in a box or a bag or a can or a dish? ____________________

____

2. How many eggs did she buy? ____________________

3. What did the children see on Pine Street? ____________________

____

4. What street was the store on? ____________________

Read this and then write the answers. Read it again if you need to.

On Monday Dick saw a red fox, a gray squirrel and a black snake in the woods. The next day he saw a brown rabbit and five brown mice in the field. He killed the fox and all the mice, but let the others live.

5. What did Dick do to the fox? ____________________

6. What colour was the rabbit? ____________________

7. Where did Dick see the fox? ____________________

Do the next page.

Read this and then write the answers. Read it again if you need to.

There was once a merchant who was very, very rich. He had six children, three boys and three girls. His daughters were all beautiful, but the youngest one was the most beautiful of all. From the time she was a small child she had been spoken of as "Beauty." As she grew older this caused a great deal of jealousy on the part of her sisters. But the young girl was not only more beautiful than they were; she was also kinder and more lovable.

8. How many brothers had Beauty? ______________________________

9. Was Beauty younger than either of her sisters? ______________________________
____

10. Was the oldest of the daughters the kindest? ______________________________

11. What word is used to tell how Beauty's sisters felt toward her as she grew older? ____

Read this and then write the answers. Read it again if you need to.

Fred lives in the country. He likes to hunt and fish, and has a gun that cost sixteen dollars. His sister Grace keeps hens and ducks, and sells the eggs. She is learning to play the piano, and goes to Miss Thomas for a lesson every Saturday. She likes music but Fred doesn't.

Do the next page.

12. What does Fred dislike? ______________________________

13. Does Grace give away the eggs which her hens lay? ______________________________

____

14. Does Fred like music less than Grace does? ______________________________

Read this and then write the answers. Read it again if you need to.

Henry and Sam are eight years old. Edward is nine. Arthur is twelve and Richard is eleven. Henry and Edward play together and like each other very much. They will not play with Sam unless they have to; for neither of them likes him at all. He feels the same way toward them. They both would be very glad to play with Arthur or Richard, but they don't often have the chance; for these boys play with boys of their own age whenever they can.

15. Which of the five boys would Arthur probably prefer to play with? __________

____

16. Name two boys who dislike each other. ______________________________

____

17. Who play with boys of their own age? ______________________________

____

18. Whom does Sam dislike? ______________________________

Do the next page.

Read this and then write the answers. Read it again if you need to.

Both before and after Christmas, Bob Adams worked harder than he did in the spring, summer or fall. Only very rarely did he reach home before eleven o'clock; and on every morning except Sunday he was up at six, dressed and done with breakfast by quarter of seven, left the house at ten minutes of seven and reached Mr. Clark's store at ten minutes of eight. In spite of the long hours and hard work, he was happy because his pay had been raised twice.

19. About how long did it take Bob to go from the house to the store? ____________

____

20. How often each week did Bob stay in bed after six? ______________________

____

21. What did Bob do between six and seven A.M. six days out of the seven? ______

____

Read this and then write the answers. Read it again if you need to.

According to the Kansas City Star, the wheat farmers of Kansas are too prosperous to trouble themselves about careful harvesting. They do not cut the fields clean. A gleaner 80 years old, after the wheat harvest in Pawnee County last year, went over the wheat-fields with a wagon, a rake, a brush, and a shovel and swept up the wheat left on the ground by the threshers. He gathered nine hundred bushels in forty days and sold it at a dollar a bushel.

Do the next page.

22. Do you think most of the wheat which the farmers grew was left on the ground by the threshers? ____________________________

23. Why are the fields not cut clear? ____________________________

____

24. Where did the gleaner spend the forty days? ______________________

____

Read this and then write the answers. Read it again if you need to.

In the spring, men begin to work the soil. Those who have small gardens break up the ground with such tools as spades and forks. Those who live on farms turn the soil over with plows drawn by horses. In these ways the soil is loosened. After the planting is done, the weeds must be killed and the soil must be loosened. In places where little rain falls, other ways of watering the plants must be found.

25. When must the soil be loosened a second time? ____________________

____

26. Write two words telling two activities of the farmer in the spring. _______

____

Do the next page.

Read this and then write the answers. Read it again if you need to.

For nearly thirty years "Lewis Carroll" was a lecturer on mathematics at Oxford. He studied divinity and occasionally preached, but his shy and retiring nature, together with a tendency to stammer, kept him from the regular ministry. He gave many lectures to audiences made up mainly of children. These lectures were of various sorts, but consisted principally of narratives from his books illustrated by lantern pictures. He invented a number of mathematical games.

27. What two causes prevented Lewis Carroll from being a regular clergyman in charge of a church? ____

____

28. What sort of lectures did he give when his audiences were not children? ____

____

____

29. What important work has he done beside lecturing to adults and children, preaching, and inventing games? ________________________________

____

____

Do the next page.

Read this and then write the answers. Read it again if you need to.

COLERIDGE

I see thee pine like her in golden story
Who, in her prison, woke and saw, one day,
The gates thrown open -- saw the sunbeams play
With only a web 'tween her and summer's glory;
Who, when the web -- so frail, so transitory,
It broke before her breath -- had fallen away,
Saw other webs and others rise for aye,
Which kept her prisoned till her hair was hoary.
Those songs half-sung that yet were all divine --
That woke Romance the queen to reign afresh
Had been but preludes from that lyre of thine,
Could thy rare spirit's wings have pierced the mesh
Spun by the wizard who compels the flesh,
But lets the poet see how heav'n can shine.

30. Who acted like a spider? ______________________________

31. Who or what is compared with the woman? ______________________________

____

32. Copy the first word of the line which implies there had not been a continuous stream of like songs ______________________________

33. Complete the following with one word only

"Those songs" really means those ______________________________

Do the next page.

Read this and then write the answers. Read it again if you need to.

There are two methods by which one might make himself acquainted with anything made up of related parts; as, for example, a watch. He might take the watch apart, piece by piece, and while doing so study the details of its structure and the relation of its parts one to another. An operation like this, which begins with the whole and descends to the parts which compose the whole, is called analysis. The word means a taking apart or separating. Or he might begin with the parts, and, after some experiment and study get an excellent knowledge of the watch by putting its parts properly together. An operation of this kind is called synthesis.

34. What two words are used in the paragraph to refer to both analysis and synthesis?

____

35. Which two of these words have a meaning similar to analysis?

Draw a line under them.

method
related
separation
experiment
operation
resolution
assemble
dissemble

Appendix B2

Form A

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PART I

TEST 3 -- PARAGRAPH READING

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name .... Boy or Girl ....
IN CAPITALS LAST FIRST

Age last Birthday .... Birthday ....
MONTH DATE YEAR

Grade .... Name of School ....

Province .... City, Town or Municipality ....

Date .... Name of Teacher ....

PARAGRAPH READING

DIRECTIONS: This is a reading examination. It consists of a number of paragraphs followed b questions to test how well you understand them. For each question five possible answers are given. I each case you are to decide which of these answers is correct and write its number in the brackets at th right of that question.

Read the sample paragraph given below, and notice how the questions are answered. In the fir question the second answer is the correct one, so 2 is written in the brackets. In the second question, th fifth answer is correct, so 5 is written in the brackets.

SAMPLE:

The typhoon is the name commonly given in the Chinese Seas to those storms which are elsewhere called hurricanes and tornadoes. The wind within the storm area rotates with frightful energy, so that in the same storm the wind may be north at one place and south at another place a few miles away.

1. The typhoon is the same as a (1) **squall** (2) **tornado** (3) **monsoon** (4) **blizzard** (5) **chinook**......(2)

2. The wind in a typhoon blows (1) **always north** (2) **vertically** (3) **always east** (4) **straight ahead** (5) **in a circle**......(5)

Now go ahead and do the rest yourself in the same way. As soon as you have read a paragraph, answe the questions following it. Work as rapidly as possible.

I had a long way to walk through the roughly-paved, mean streets of the little seaport, dimly lighted as they were by the miserable oil-lamps that swung in the boisterous wind. The shutters were up everywhere, though now and then a gush of light would stream from between their chinks, and I could hear voices, chatting over the cozy fires and supper-tables. The public-houses that I passed were not empty, for there was a crimson glow visible through their red curtains and voices were heard singing or disputing in sea slang.

1. The town was (1) **by the ocean** (2) **in the mountains** (3) **by a lake** (4) **in farming country** (5) **on a river**......( )

2. The night was (1) **quiet** (2) **clear** (3) **windy** (4) **warm** (5) **wet**......( )

3. Public-houses are (1) **schools** (2) **dance-halls** (3) **churches** (4) **tenements** (5) **inns**...( )

Continue on page 2

Feudalism, which flourished in Europe during the Middle Ages, was a form of government based on the ownership of land. Under this system kings and nobles divided their estates among feudal lords who, in turn, granted sections of land to lesser tenants, and so on. Land so granted was called a fief, the grantor was called lord or liege, and the one who received it, a vassal. The vassal pledged fealty to his lord and received, in return, protection against the lawlessness of the time. Lowest in rank of the several classes of vassals were the serfs, who were bound to the soil and who passed with the land when it changed hands.

4. Fealty means the same as **(1) enmity (2) loyalty (3) hope (4) dishonesty (5) contempt**.......................................( )

5. A serf was a **(1) lord (2) feud (3) fief (4) liege (5) vassal**.......................................( )

6. Another name for lord was **(1) vassal (2) liege (3) fief (4) serf (5) tenant**.......................................( )

7. The Middle Ages were **(1) enlightened (2) democratic (3) lawless (4) scientific (5) peaceful**.......................................( )

8. A fief was **(1) a lesser tenant (2) a period in the Middle Ages (3) a piece of land (4) a form of government (5) a court of law**.......................................( )

9. We might infer from this paragraph that feudalism existed in **(1) China (2) Australia (3) Canada (4) France (5) Egypt**.......................................( )

The Coastguard service, organized originally for the sole purpose of guarding against smuggling, is now primarily devoted to life-saving and general coast defence, although its former duties remain. For the purpose of the organization of this service the coasts of Great Britain are divided into six districts, each district is sub-divided into forty-four divisions, and each division into several stations. Each district is under a captain, who has a cruiser and smaller boats at his disposal; lieutenants and petty officers are in charge of divisions and stations respectively. The men are, as a rule, recruited from the navy, and have usually seen nine or ten years of service. They receive relatively high pay, besides quarters and other privileges.

10. Recruits come to the Coastguard service from **(1) the army (2) training schools (3) shops and farms (4) the navy (5) the university**.......................................( )

11. The Coastguard has always had as one of its duties **(1) life-saving (2) erecting light-houses (3) coast defence (4) policing harbours (5) preventing smuggling**.......................................( )

12. We would infer that the number of coastguard cruisers in Great Britain is **(1) 1 (2) 6 (3) 44 (4) 50 (5) 264**.......................................( )

13. Lieutenants have charge of **(1) stations (2) coasts (3) divisions (4) cruisers (5) districts**.......................................( )

14. As used here, "quarters" refer to **(1) lodgings (2) pin money (3) holidays (4) bonuses (5) pensions**.......................................( )

Continue on page 3

"Your letter conveys a picture of married happiness which indeed raises my envy. In spite of my single life and wandering habits, which are the result of necessity rather than desire, I am really quite a domestic sort of fellow, and I cannot hear of my old cronies snugly nestled down with good wives and fine children round them, but I feel for the moment desolate and forlorn."

15. The writer of this paragraph is probably writing to (1) **a boyhood friend** (2) **his wife**
(3) **his sweetheart** (4) **a comparative stranger** (5) **his father**....................( )

16. The theme of the paragraph is the writer's longing (1) **to travel** (2) **for freedom**
(3) **to return home** (4) **for married life** (5) **for wealth**..........................( )

Algae, plants of the simplest structure, live in water, both sea and fresh, or clothe damp surfaces, such as rocks, walls or the bark of trees. Unlike fungi, all contain green colouring matter and require light for their existence. Some consist of a single cell of very minute size, others of long slender filaments, while others again reach gigantic proportions. They have no true roots, but seem to absorb nourishment at all parts of their surface, from the water, or damp air with which they are surrounded.

17. Algae, for their existence, seem to require (1) **rich soil** (2) **protection from light**
(3) **extreme warmth** (4) **a great deal of moisture** (5) **constant care**.............( )

18 Fungi (1) **all reach gigantic proportions** (2) **contain no green colouring matter**
(3) **are very complex in structure** (4) **require light for their existence**
(5) **have no true roots** ....( )

The new history text is a good and interesting outline by a practical teacher, though some of the simplifications are, as the author himself recognizes, open to question. The numerous illustrations are of great value and of more unusual type than in many textbooks, while the abstracts are from original sources and the exercises at the end of each chapter will be welcomed by busy teachers. The book is intended for use in the middle form.

19 The abstracts in the new text are (1) **carefully copied** (2) **of unusual type**
(3) **more numerous than usual** (4) **less numerous than usual**
(5) **from original sources** ....( )

20. The illustrations are (1) **of the usual type** (2) **from original sources** (3) **numerous**
(4) **open to question** (5) **at the ends of the chapters**............................( )

End of Reading Test. Look over your work.

Pupil's score (number right) ————

Form B

THE DOMINION TESTS

GROUP ACHIEVEMENT TEST

PARAGRAPH READING

INSTRUCTIONS TO PUPILS

Listen closely to what you are told to do and work quickly and carefully.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Name ....Boy or Girl ....

Age ....Birthday ....
MONTH DATE YEAR

Grade ....Name of School ....

Date .... Name of Teacher ....

Paragraph Reading Total Score ________________

TEST 3. PARAGRAPH READING

DIRECTIONS: This is a reading examination. It consists of a number of paragraphs followed by questions to test how well you understand them. For each question five possible answers are given. In each case you are to decide which of these answers is correct and write its number in the brackets at the right of that question.

Read the sample paragraph given below, and notice how the questions are answered. In the first question the second answer is the correct one, so 2 is written in the brackets. In the second question, the fifth answer is correct, so 5 is written in the brackets.

SAMPLE:

The typhoon is the name commonly given in the Chinese Seas to those storms which are elsewhere called hurricanes and tornadoes. The wind within the storm area rotates with frightful energy, so that in the same storm the wind may be north at one place and south at another place a few miles away.

1. The typhoon is the same as a (1) **squall** (2) **tornado** (3) **monsoon** (4) **blizzard** (5) **chinook**..(2)

2. The wind in a typhoon blows (1) **always north** (2) **vertically** (3) **always east** (4) **straight ahead** (5) **in a circle**..(5)

Now go ahead and do the rest yourself in the same way. As soon as you have read a paragraph, answer the questions following it. Work as rapidly as possible.

The beehive usually contains three kinds of occupants, a single perfect female or "queen" bee, the drones or males, and the workers or imperfect females. The working bees, in addition to visiting flowers in search of honey and pollen, attend on the queen, make and prepare the cells for the reception of the eggs, and feed the larvae and young bees.

1. The workers are (1) **perfect males** (2) **imperfect females** (3) **larvae** (4) **imperfect male** (5) **perfect females**..( )

2. The larvae are fed by the (1) **young bees** (2) **drones** (3) **queens** (4) **males** (5) **workers**.( )

3. Usually, the number of "queens" in a beehive is (1) **one** (2) **several** (3) **three** (4) **none** (5) **many**..( )

Continue on page 2

It was on one of the lovely mornings of our ever lovely autumn, so early that the sun had scarcely touched the tops of the still verdant forest, that Silas Ashburn and his eldest son sallied forth for a day's chopping on the newly-purchased land of a rich settler, who had been but a few months among us. The tall form of the father, lean and gaunt as the very image of Famine, derived little grace from the rags that streamed from the elbows of his almost sleeveless coat, or flapped round the tops of his heavy boots. Poor Joe's costume showed, if possible, a still greater need of the aid of that useful implement, the needle. His mother is one who thinks little of the ancient proverb that commends the stitch in time; and the clothing under her care sometimes falls to pieces, seam by seam.

4. **The time of day was probably about (1) 1 a.m. (2) 7 a.m. (3) 11 a.m. (4) 6 p.m. (5) 10 p.m.**..( )

5. **The father was (1) thin (2) happy (3) strong (4) short (5) rich**..................( )

6. **Joe's mother was (1) neat (2) energetic (3) lazy (4) thrifty (5) ugly**................( )

7. **The month may have been (1) February (2) April (3) July (4) October (5) January**..( )

A light shone from the window of the hut, and swiftly they advanced towards it. Passing through the wall of mud and stone, they found a cheerful company assembled round a glowing fire, an old, old man and woman, with their children and their children's children, and another generation beyond that, all decked out gaily in their holiday attire. The old man, in a voice that seldom rose above the howling of the wind upon the barren waste, was singing them a Christmas song; it had been a very old song when he was a boy; and from time to time they all joined in the chorus. So surely as they raised their voices, the old man got quite blithe and loud; and so surely as they stopped, his vigour sank again.

8. The words "another generation" refer to the old people's **(1) grandchildren (2) parents (3) children (4) great-grandchildren (5) cousins**................................( )

9. The country in which the hut was situated was **(1) bare (2) fertile (3) mountainous (4) wooded (5) beautiful**..( )

10. When the others joined in, the old man grew **(1) quiet (2) angry (3) loud (4) impatient (5) sober**...( )

Continue on page 3

Amphibia are vertebrate or backboned animals which breathe by means of gills during the early period of their development, but possess lungs in the adult stage. They can be grouped into four orders or classes. In the first, of which the blind-worms of the Tropics are an example, the body is snake-like and without legs: in the second the tail of the embryo persists throughout life, as seen in the newt; the third, which includes frogs and toads, has neither tail nor gills when grown to maturity; the fourth order is extinct. In all existing species the young are produced from eggs, and all at first possess the form and organs of fish.

11. We would infer from this paragraph that blind-worms flourish best **(1) in dry climates (2) where it is warm (3) at high altitudes (4) in cold regions (5) in mountainous country**..........( )

12. The term "extinct" is applied to classes of animals which **(1) occur in great abundance (2) are snake-like in form (3) are no longer in existence (4) breathe by gills (5) eat only vegetable matter**..........( )

13. Amphibia spend their early life **(1) in holes in the ground (2) on plants (3) under stones and logs (4) clinging to other animals (5) in water**..........( )

14. The newt **(1) is extinct (2) belongs to the third order (3) is legless (4) has a tail (5) is blind**..........( )

15. All vertebrates have **(1) legs (2) hair (3) gills (4) tails (5) bones**..........( )

But neither climate nor poverty, neither study nor the sorrows of a home-sick exile, could tame the desperate audacity of Clive's spirit. He behaved to his official superiors as he had behaved to his schoolmasters, and was several times in danger of losing his situation.

16. Clive was **(1) dishonest (2) lazy (3) weak (4) poor (5) meek**..........( )

17. At school Clive had been **(1) obedient (2) quiet (3) unruly (4) patient (5) gentle**..( )

18. Audacity is **(1) strength (2) impatience (3) vanity (4) despair (5) boldness**..... ( )

The personal strength of the outlaw was not equal to his activity; but his wit so far excelled his might that he never found use for the strength which he had—so well did he form his plans and work out his stratagems.

19. The outlaw owed his success mainly to his **(1) strength (2) intelligence (3) luck (4) activity (5) might**..........( )

20. Stratagems are **(1) schemes (2) problems (3) ambitions (4) solutions (5) falsehoods**..........( )

End of Reading Test. Look over your work.

Pupil's score (number right).―――――

Appendix B4

Administration, Test #1, 90 Common Names

NOTE TO TEACHERS

RE: VOCABULARY TEST -- JANUARY 1976

1. Once all of the pupils are assembled, please distribute the test forms to them. Writing and printing of answers in pen or pencil is acceptable. Instruct the children to enter their name, age, grade, school, and the date in the spaces provided. Every child should write or print his birthdate (month, day and year -- e.g., April 12, 1960) in the top right-hand corner of the test cover.

> SPECIAL NOTE: All grade 7 and 8 students should enter the name of their Grade 6 school below the name of their present school.

2. When all children have completed this task, the following instructions should be read to them.

"This is a test to find out how many words you know. You are to put the answer to each question in the space provided below each question.

Follow the paper as I read:

> *Parts of the Body.* What do we call the:
> 1. Part of the leg above the knee?

What is the answer? (answer -- thigh)

Write the word thigh on the line below question number 1. (Teacher writes it on the blackboard: #1 -- thigh.)

Look at question No. 6:

> *Trades.* What do we call a man who:
> 6. Makes clothes?

What is the answer? (tailor) Write the word tailor in the space provided. (Teacher writes #6 -- tailor on the blackboard.)

(Teacher: Do the same with *Rooms* No. 11 -- attic and
Dwellings No. 28 -- garage.)

Now, go back to Question No. 2. You are to do all the questions in this test in the same way. If you have trouble with spelling, do the best you can. If you cannot answer a question, skip it and go on the next.

Do your best because we want you to do as well as the other Grade (5, 6, 7, or 8) classes in the city. You will have 50 minutes to do the test. Ready, begin."

3. Due to the variety of changes in the English language and its instruction in the elementary schools, it is expected that the children will not be familiar with a number of the names in the test. In order to determine the extent of this change, the test must be administered in its original form.

4. Please record the School, Date, Grade and Number of Pupils tested on the front of the return envelopes.

All tests administered in your school will be picked up on *January 30, 1976*, unless special arrangements have been made.

Appendix B5

STUDENT'S NAME ______________________ DATE ______________

AGE ______________________ GRADE ______________________

SCHOOL ____

VOCABULARY TEST #1

90 COMMON NAMES

PARTS OF THE BODY - What do we call the:

1. Part of the leg above the knee?

____

2. Middle part of the eye?

____

3. Parts of the nose through which we breathe?

____

4. Front part of the leg below the knee?

____

5. Bottom part of the back of the neck?

____

TRADES - What do we call a man who:

6. Makes clothes?

____

7. Mends gas-pipes and water-pipes?

____

8. Makes cakes and pies and sells them?

9. Is paid to drive a private motor-car?

ROOMS - What do we call a:

10. Reading-room with a great many books in it?

11. Room at the top of a house just under the roof?

12. Room where clerks and typists work?

RECEPTACLES - What do we call a:

13. Metal or wooden box for keeping small quantities of the tea in?

14. Glass for drinking water out of?

15. Large open-work packing-case for oranges, eggs, etc.?

16. Wooden or metal box for keeping small supplies of coal in?

TOOLS - What do we call the tool used for:

17. Driving nails into wood?

18. Grasping and holding lumps of coal?

19. Cutting grass and clipping hedges?

20. Levelling the surface of pieces of wood?

21. What do we call a large blunt needle for drawing tape, ribbons, etc, through hems?

SHOPKEEPERS - What do we call a man who keeps a store for the sale of:

22. Such things as currants, rice, sugar, tea, coffee, and cocoa?

23. Cigarettes, matches, pipes and smokers' outfits?

24. Such articles as microscopes, spectacles, lenses and opera-glasses?

25. Writing-paper, pens, pencils, ink, etc., but not newspapers or stamps?

DWELLINGS, ETC. - What do we call a:

26. Wooden shelter made for a dog to sleep in?

27. Place made for a horse to feed and sleep in?

28. Place where motor-cars may be kept?

29. Large inn specially fitted up for travellers and holiday makers?

30. A one-storey dwelling-house?

OPENINGS - What do we call:

31. A hole accidentally made in a bicycle-tire?

____

32. Openings made for pennies in an automatic gum machine?

____

33. The passage for smoke running up through a chimney?

____

34. The holes in a net or sieve?

____

35. The small holes in a sheet of postage-stamps separating one stamp from another?

____

MOTOR-CARS - What do we call the:

36. Part which the driver uses when he wants to stop?

____

37. Big bright lights seen at the front of a motor-car at night?

____

38. Frame of glass which the driver looks through when he is driving?

____

39. Rubber covers fitted around the rims of the wheels?

____

40. Bars which run out from the middle of the wheel to the rim?

____

HEAD-DRESS - What do we call:

41. The "hat" which a policeman wears?

____

42. The covering which a baby wears on her head and tied under her chin?

____

43. The covering which a woman sometimes wears on her head when doing housework?

44. The part of a cloak made to turn up, when required, over the head?

BEDS AND BEDDING - What do we call the:

45. Part of the bed which is usually stuffed with felt or kapok?

46. Little wheels fitted to the legs of a bed to make it easy to move?

47. Framework of a bed (i.e., the part left after all the bedclothes have been taken off)?

ASSEMBLIES - What do we call a number of people:

48. Gathered together to hear a concert?

49. Gathered together to go for a picnic or outing?

50. Gathered together at a church service?

51. Selected to sing together?

52. Gathered together in a disorderly crowd?

FLOWERS AND PLANTS - What do we call the:

53. Brightly coloured parts of a flower?

54. Lines or ribs on the leaves of a plant?

55. Powder or dust found in flowers?

56. Slender twining parts which some plants have to help them to climb?

57. A small woody tree-like plant or bush?

HOUSEHOLD FITTINGS - What do we call the:

58. Floor of the fireplace in front of the fire?

59. Fittings on which doors hang and swing?

60. Upright supports for the handrail in a staircase or balcony?

61. Sets of steps in a staircase?

62. Sloping beams in a roof?

BOOKS - What do we call the:

63. Pictures and drawings in a book?

64. Sheets of paper which we turn over, one by one, in reading?

65. Lettering in a book?

66. Part which holds the cover to the rest of the book?

67. Page at the beginning of a book which tells us in large letters what the book is called?

FRUITS AND SEEDS - What do we call the:

68. Middle part of an apple with the seeds in it?

69. Skin of an apple or pear?

70. Inner part of a nut or plum stone?

71. Fleshy part of fruits?

72. Green cases or shells in which peas and beans grow?

PARTS OF A HOUSE - What do we call:

73. The entrance-passage sometimes found in a house?

74. A window set in a roof or ceiling?

75. The under part of a roof overlapping the walls?

76. The level flooring at the top of a staircase?

77. A window which stands out in front of the rest of the house?

MATERIALS - What do we call the:

78. Liquid metal often found in thermometers?

79. Metal which is most commonly used for wedding rings?

80. Metal from which the best knife-blades are made?

81. Reddish-brown metal used for making wires which have to carry electric currents?

DIVISIONS - What do we call the:

82. Separate divisions of the glass in a window?

83. Main divisions or sections of a story-book?

84. Portions of a bottle of medicine (e.g., teaspoonfuls) as we take them?

85. Parts of a sum of money which have to be paid regularly?

MEASURING AND TESTING - What do we call the instruments used for:

86. Measuring how hot something is?

87. Testing whether a table surface is quite flat?

88. Measuring the amount of gas or electricity which has been used?

89. Measuring how heavy things are?

90. What do we call a narrow band of material marked in feet and inches used for measuring cloth, etc.?

Appendix B6

Administration, Test #1, 90 Common Names

NOTE TO TEACHERS

RE: REVISED VOCABULARY TEST -- APRIL - MAY, 1976

Once all of the pupils are assembled, please distribute the test forms to them. Printing of answers in pen or pencil on the test paper is acceptable. Instruct the children to enter their name, age, grade, school, and the date in the spaces provided. Every child should write or print his birthdate (month, day and year -- e.g., April 12, 1960) in the top right-hand corner of the test cover.

> SPECIAL NOTE: All grade 7 and 8 students
> should enter the name of their Grade 6 school
> below the name of their present school.

When all children have completed this task, the following instructions should be read to them.

"This is a test to find out how many words you know. You are to put the most appropriate answer to each question in the space provided below each question.

Follow the paper as I read:

> *Parts of the Body.* What do we call the:
> 1. Part of the leg above the knee?

What is the answer? (Answer: thigh)

Write the word thigh on the line below question number 1. (Teacher writes it on the blackboard: #1 -- thigh.)

Look at question No. 6:

> *Trades.* What do we call a man who:
> 6. Makes clothes?

What is the answer? (Answer: tailor)

Write the word tailor in the space provided. (Teacher writes #6 -- tailor on the blackboard.)

(Teacher: Do the same with: *Rooms.* No. 11 -- basement and
Dwellings. No. 28 -- garage.)

Now, go back to Question No. 2. You are to do all the questions in this test in the same way. If you have trouble with spelling, do the best you can. If you cannot answer a question, skip it and go on to the next.

Read each question very carefully. (Only one answer will be accepted for each question.)

Do your best because we want you to do as well as the other Grade (5, 6, 7, or 8) classes in the city. You will have 50 minutes to do the test. Ready. Begin."

TEACHER:

Please record the School, Date, Grade and Number of Pupils tested on the front of the return envelopes.

All tests administered in your school will be picked up on May 24, 1976 unless special arrangements have been made.

We realize the inconveniences associated with this testing programme and sincerely appreciate your cooperation!

PLEASE PRINT ALL ANSWERS.

STUDENT'S NAME ______________________ DATE ____________

AGE ____________________ GRADE ______________________

SCHOOL ____

VOCABULARY TEST #1

(NIAGARA TEST SERIES)

90 COMMON NAMES

PARTS OF THE BODY - What do we call the:

1. Part of the leg above the knee?

____

2. Middle part of the eye?

____

3. Parts of the nose through which we breathe?

____

4. Back part of the leg below the knee?

____

5. Bottom part of the ear on which an earring is worn?

____

TRADES - What do we call a person who:

6. Makes clothes?

____

7. Repairs gas-pipes and water pipes?

____

8. Makes cakes and pies and sells them?

____

9. Is paid to drive a private automobile?

....

ROOMS - What do we call a:

10. Reading-room with a great many books in it?

....

11. Room of a house that is partly or totally underground?

....

12. Room where clerks and typists work?

....

RECEPTACLES - What do we call a:

13. Bag for carrying belongings on one's back?

....

14. Heavy cup used without a saucer, out of which you drink hot chocolate or coffee?

....

15. Container holding one dozen eggs?

....

16. Container used to hold laundry?

....

TOOLS - What do we call the tool used for:

17. Driving nails into wood?

....

18. Tightening bolts?

....

19. Trimming lawn edges and clipping hedges?

....

20. Levelling the surface of pieces of wood?

....

21. Flattening out pie dough?

....

SHOPKEEPERS - What do we call a person who keeps a store for the sale of:

22. Such things as fresh flowers, flower arrangements and flower containers?

....

23. Medicines, health aids, cosmetics, etc.?

....

24. Such things as rings, watches, necklaces, etc.?

....

25. Writing-paper, pens, pencils, ink, etc., but not newspapers or stamps?

....

DWELLINGS, OR BUILDINGS, ETC. - What do we call a:

26. Building for watching or playing hockey games?

....

27. Place made for a horse to feed and sleep in?

....

28. Place where automobiles may be kept?

....

29. Large building specially equipped for servicing travellers and people on holiday?

....

30. Place where people can go to buy a meal?

....

OPENINGS - What do we call:

31. A hole accidentally made in a bicycle-tire?

....

32. Openings made for coins in a pay telephone or pop machine?

....

33. The hole made near the end of a sewing needle?

....

34. The very tiny openings in the skin?

....

35. The opening(s) through which water flows out of a sink?

....

AUTOMOBILES - What do we call the:

36. Part which the driver uses when he wants to stop?

....

37. Big bright lights seen at the front of an automobile at night?

....

38. Glass through which the driver looks when he is driving?

....

39. Rubber covers fitted around the rims of the wheels?

....

40. Rods which run out from the middle of the wheel to the rim?

....

HEAD-DRESS - What do we call:

41. The "hat" which a policeman wears?

....

42. The covering which a baby wears on his/her head and tied under his/her chin?

....

43. The covering which a woman sometimes wears on her head when doing housework?

....

44. The part of a cloak, jacket, or coat made to turn up, when required, over the head?

....

BEDS AND BEDDING - What do we call the:

45. Part of the bed which is usually padded with felt or foam rubber?

....

46. Stuffed object used to support the head while sleeping?

....

47. Special bed made for a baby?

....

ASSEMBLIES - What do we call a number of people:

48. Gathered together to hear a concert?

....

49. Gathered together on a regular basis to play sports?

....

50. Who work together on a ship?

....

51. Selected to sing together?

....

52. Gathered together in a disorderly crowd?

....

FLOWERS AND PLANTS - What do we call the:

53. Brightly coloured parts of a flower?

....

54. Lines or ribs on the leaves of a plant?

....

55. Powder or dust found in flowers?

....

56. Slender twining parts which some plants have to help them to climb?

....

57. Small woody tree-like plant or bush?

....

HOUSEHOLD FITTINGS - What do we call the:

58. Shield above a door or window to keep out the sun, rain, and snow?

....

59. Fittings on which doors hang and swing?

....

60. Hand support of a staircase?

....

61. Wooden or metal structure used to support the glass in a window?

....

62. Metal conductor along a roof used to transfer rainwater?

....

BOOKS - What do we call the:

63. Pictures and drawings in a book?

....

64. Sheets of paper which we turn over, one by one, in reading?

....

65. Person who wrote the book?

....

66. Section, usually at the back of a book, which states the pages where certain information may be found?

....

67. Page at the beginning of a book which tells us in large letters what the book is called?

....

FRUITS AND SEEDS - What do we call the:

68. Middle part of an apple with the seeds in it?

....

69. Skin of a potato or cucumber?

....

70. Hard seed found inside a cherry or peach?

....

71. Fleshy part of a tomato or other fruit?

....

72. Green cases or shells in which peas and beans grow?

....

PARTS OF A HOUSE - What do we call:

73. The cabinets in a kitchen used for storage of glassware and dishes?

....

74. A place where clothes are hung or stored?

....

75. The individual pieces of protective covering on a roof?

....

76. The flat part below the bottom of a door or window?

....

77. The different levels in a house or office building?

....

MATERIALS - What do we call the:

78. Liquid metal often found in thermometers?

....

79. Metal which is most commonly used for wedding rings?

....

80. Metal from which the best knife-blades are made?

....

81. Reddish-brown metal used for making wires which have to carry electric currents?

....

DIVISIONS - What do we call the:

82. Separate pieces of the glass in a window?

....

83. Main sections of a story-book?

....

84. Individual pieces of bread cut from a loaf?

....

85. Three parts or divisions of a hockey game?

....

MEASURING AND TESTING - What do we call the instruments used:

86. For measuring how hot something is?

....

87. For measuring and recording the amount of electricity used in a home or business?

....

88. By police for determining the speed of passing cars?

....

89. For measuring how heavy things are?

....

90. What do we call a narrow band of material marked in mm, cm, and m, used for measuring cloth, etc.?

....